



STIC Search Report

Biotech-Chem Library

STIC Database Tracking Number: 122947

TO: Frank Prats
Location: REM/3A41/3E71
Art Unit: 1651
Thursday, May 27, 2004

Case Serial Number: 09/807146

From: Mary Jane Ruhl
Location: Biotech-Chem Library
Remsen 1-B55
Phone: 571-272-2524

maryjane.ruhl@uspto.gov

Search Notes

Examiner Prats,

Here are the results for your recent search request.

Please feel free to contact me if you have any questions about these results.

Thank you for using STIC services. We appreciate the opportunity to serve you.

Sincerely,

Mary Jane Ruhl
Technical Information Specialist
STIC
CM-1, Rm. 6-A-06
605-1155

Query Match 100.0%; Score 3408; DB 3; Length 636;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 636; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 MLTPQOYGLIIQYKTRIIDITYPEQAGIEKSEDMQFQSRMDTHPPKLNELDSYVG 60
DB 1 MLTPQOYGLIIQYKTRIIDITYPEQAGIEKSEDMQFQSRMDTHPPKLNELDSYVG 60
QY 61 NNEALLPMLLEMLLAQAMQSYQORNSSLKDIDIARENNPDWILSNKQVGVGYVDFAGDL 120
DB 61 NNEALLPMLLEMLLAQAMQSYQORNSSLKDIDIARENNPDWILSNKQVGVGYVDFAGDL 120
QY 121 KGLKDKIYPFOELGLTYLHMLPLFKCPBGKSDGYAVSSYRDVNPALGTIGDLREVIATL 180
DB 121 KGLKDKIYPFOELGLTYLHMLPLFKCPBGKSDGYAVSSYRDVNPALGTIGDLREVIATL 180
QY 181 HEAGISAVVDPIFNHTSNEHEMAQRCAGDPLFDFNYIIFPDRMPDQYDRLREIFPDQ 240
DB 181 HEAGISAVVDPIFNHTSNEHEMAQRCAGDPLFDFNYIIFPDRMPDQYDRLREIFPDQ 240
QY 241 HPGGSQLEDGRWWTTFNSFQMDLNTSNPVFRAMAGEMFLANLGVDIILMDAVATIW 300
DB 241 HPGGSQLEDGRWWTTFNSFQMDLNTSNPVFRAMAGEMFLANLGVDIILMDAVATIW 300
QY 301 KQMGTCENLPQAHALIRAFNAVMRIAPAVFEKSEALVHPDQVVOYIGDECOIGYNPL 360
DB 301 KQMGTCENLPQAHALIRAFNAVMRIAPAVFEKSEALVHPDQVVOYIGDECOIGYNPL 360
QY 361 QMALIMNTLATREVNLLHQALTYRHNLPEHTAMVNYVSHDDIGTFPDEDAAYIGISGY 420
DB 361 QMALIMNTLATREVNLLHQALTYRHNLPEHTAMVNYVSHDDIGTFPDEDAAYIGISGY 420
QY 421 DHRQFLNRFVNRFDGSPARQVFPQYNPSTGDCRSYGAALVGLAODPHAVDRIKLLY 480
DB 421 DHRQFLNRFVNRFDGSPARQVFPQYNPSTGDCRSYGAALVGLAODPHAVDRIKLLY 480
QY 481 STALSTGGLPIIYIGDEVGTINDDWSQDSNKSDDSRMAHPRRYNEALYAOENDESTAAG 540
DB 481 STALSTGGLPIIYIGDEVGTINDDWSQDSNKSDDSRMAHPRRYNEALYAOENDESTAAG 540
QY 541 QIYQGLRMIIVAVROSNPRFDGRLTYTNTNNKHIIIGYIRNNALLAFGNSEYPTVTAHT 600
DB 541 QIYQGLRMIIVAVROSNPRFDGRLTYTNTNNKHIIIGYIRNNALLAFGNSEYPTVTAHT 600
QY 601 LOAMPFKAHDLIGKTVSLNODLTLPYQVWMLLEIA 636
DB 601 LOAMPFKAHDLIGKTVSLNODLTLPYQVWMLLEIA 636
  
```

RESULT 2
 AAM90979
 ID AAM90979 standard; protein; 636 AA.
 AC AAM90979;
 XX
 DT 21-JUL-2000 (first entry)
 XX
 DE N. polysacchara amylosucrase protein.
 XX
 KM Alpha-1,6-branched alpha-1,4-glucan; branching enzyme; binder; carrier;
 KM transgenic plant; flavor; perfume; packaging material; papermaking;
 KM ultra-violet light absorber; starch; textile; wetting agent;
 KM amylosucrase.
 XX
 OS Neisseria polysacchara.
 XX
 PN WO200022140-A1.
 XX
 PD 20-APR-2000.
 XX
 PF 08-OCT-1999; 99MO-EP007562.
 XX
 PR 09-OCT-1998; 98DE-01046635.

PR 27-MAY-1999; 99DE-01024342.
 XX
 XX (PLAN-) PLANTTEC BIOTECHNOLOGIE GMBH.
 PA (PLAC) MAX PLANCK GES FOERDERUNG WISSENSCHAFTEN.
 XX
 PI Buettcher V, Quanz M;
 XX
 XX MPI; 2000-317992/27.
 DR N-PSDB; AAA11732.
 XX
 PT New nucleic acid encoding a branching enzyme, useful for in vitro
 PT synthesis of branched glucans and to prepare transgenic plants producing
 PT modified starch.
 PS Disclosure; Page 102-104; 115pp; German.

This invention describes a novel nucleic acid (I) isolated from *Neisseria* which encodes a branching enzyme (II). (I) is used for recombinant production of (II) subsequently used in the in vitro production of alpha-1,6-branched alpha-1,4-glucans. It is also used to prepare transgenic plants that produce starches with modified properties. (III) are used as binders for tablets, carriers for pharmaceuticals, flavors and perfumes and powdered additives, packaging materials, ultra-violet light absorbers in sunscreens and also for any of the usual applications of starch in foods, papermaking, as textile size, in soil stabilization, as wetting agent for agricultural chemicals, as polymer additives etc. Fragments of (I) are useful as PCR primers and antisense molecules or ribozymes for inhibiting expression of (I), and the regulatory region of (II) can be used to control expression of heterologous sequences in host cells. (I) provides an inexpensive method for producing alpha-1,6-branched alpha-1,4-glucans (III), producing products that can be tailored for particular applications, particularly by controlling the degree of branching. Starch from transgenic plants has increased gel strength; reduced phosphate content; reduced peak viscosity; lower pasting temperature and granule size and/or altered sidechain distribution. This sequence represents an amylosucrase isolated from *Neisseria polysacchara* which is described in the method of the invention

Sequence 636 AA;

```

SQ
  
```

Query Match 100.0%; Score 3408; DB 3; Length 636;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 636; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 MLTPQOYGLIIQYKTRIIDITYPEQAGIEKSEDMQFQSRMDTHPPKLNELDSYVG 60
DB 1 MLTPQOYGLIIQYKTRIIDITYPEQAGIEKSEDMQFQSRMDTHPPKLNELDSYVG 60
QY 61 NNEALLPMLLEMLLAQAMQSYQORNSSLKDIDIARENNPDWILSNKQVGVGYVDFAGDL 120
DB 61 NNEALLPMLLEMLLAQAMQSYQORNSSLKDIDIARENNPDWILSNKQVGVGYVDFAGDL 120
QY 121 KGLKDKIYPFOELGLTYLHMLPLFKCPBGKSDGYAVSSYRDVNPALGTIGDLREVIATL 180
DB 121 KGLKDKIYPFOELGLTYLHMLPLFKCPBGKSDGYAVSSYRDVNPALGTIGDLREVIATL 180
QY 181 HEAGISAVVDPIFNHTSNEHEMAQRCAGDPLFDFNYIIFPDRMPDQYDRLREIFPDQ 240
DB 181 HEAGISAVVDPIFNHTSNEHEMAQRCAGDPLFDFNYIIFPDRMPDQYDRLREIFPDQ 240
QY 241 HPGGSQLEDGRWWTTFNSFQMDLNTSNPVFRAMAGEMFLANLGVDIILMDAVATIW 300
DB 241 HPGGSQLEDGRWWTTFNSFQMDLNTSNPVFRAMAGEMFLANLGVDIILMDAVATIW 300
QY 301 KQMGTCENLPQAHALIRAFNAVMRIAPAVFEKSEALVHPDQVVOYIGDECOIGYNPL 360
DB 301 KQMGTCENLPQAHALIRAFNAVMRIAPAVFEKSEALVHPDQVVOYIGDECOIGYNPL 360
QY 361 QMALIMNTLATREVNLLHQALTYRHNLPEHTAMVNYVSHDDIGTFPDEDAAYIGISGY 420
DB 361 QMALIMNTLATREVNLLHQALTYRHNLPEHTAMVNYVSHDDIGTFPDEDAAYIGISGY 420
QY 421 DHRQFLNRFVNRFDGSPARQVFPQYNPSTGDCRSYGAALVGLAODPHAVDRIKLLY 480
  
```

```

Dh 421 DHRGLNFFVNRFDGSPFARGVFPQYNSTGDCRVSTAAALVGLAQDDPHAVDRKILLY 480
Qy 481 STALSTGGLPLIYLDEVTGLNDNDWSDQNSKSDSRMAHRRRYNEALYAQNDSTAG 540
Dh 481 STALSTGGLPLIYLDEVTGLNDNDWSDQNSKSDSRMAHRRRYNEALYAQNDSTAG 540
Qy 541 QIYQGLRHMLAVROSNPRFDGRLVTFNTNNKHIIIGYIRNNALLAFGNFSEYPTVTAAH 600
Dh 541 QIYQGLRHMLAVROSNPRFDGRLVTFNTNNKHIIIGYIRNNALLAFGNFSEYPTVTAAH 600
Qy 601 LQAMPFKADLIGKTVSLNODLTLOPYQVWMLA 636
Dh 601 LQAMPFKADLIGKTVSLNODLTLOPYQVWMLA 636

```

RESULT 3

```

AA87831
ID AA87831 standard; protein; 636 AA.
XX
AC AA87831;
XX
DT 01-SEP-2000 (first entry)
XX
DE N. polysacchara amylosucrase protein.
XX
KW Polylglucan; amylosucrase; active ingredient carrier; cosmetic;
KM agrochemical; food; food additive; flavor; perfume.
XX
OS Neisseria polysacchara.
XX
PV WO200022155-A2.
XX
PD 20-APR-2000.
XX
PF 07-OCT-1999; 99WO-EP007518.
XX
PR 09-OCT-1998; 98DE-01046492.
XX
PA (AVENTIS RES & TECHNOLOGIES GMBH & CO KG.
PI Gallert K, Bengs H, Simandi C;
XX
DR WPI; 2000-318003/27.
XX
PT New polylglucan and derivatives, useful e.g. as carriers for
XX pharaceuticals, agrochemicals and cosmetics, prepared using polylglucan
XX sucrose and an additional biogenic compound.
XX
PS Claim 3; Page 24-26; 26pp; German.
XX

```

This invention describes a novel polylglucan and its derivatives (A) which is produced using a polylglucan sucrose or amylosucrase in the presence of at least one other biogenic material (I). (A) are used; (i) as carriers for active ingredients especially for pharmaceutical, cosmetic and agrochemical compounds; (ii) as food or food additives; and (iii) as carriers for flavors or perfumes. They may also be used as starting materials for further chemical modification. The activity of amylosucrase is not adversely affected by presence of (I), particularly other enzymes, and the combination of enzymes allows a very wide range of modified products to be prepared. In vitro reaction facilitates production of reproducible products of consistent quality and value, with properties (e.g. molecular weight, primary structure, polydispersity etc.) tailored for specific applications. This sequence represents a Neisseria polysacchara amylosucrase which is described in the method of the invention

Sequence 636 AA;

Query Match 100.0%; Score 3408; DB 3; Length 636;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 636; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Qy 1 MLTPQVGLILQYLKTRILDIYTPBOAGIEKSEDMQFSRMDTHFPKLMNELDSYVG 60
Dh 1 MLTPQVGLILQYLKTRILDIYTPBOAGIEKSEDMQFSRMDTHFPKLMNELDSYVG 60
Qy 61 NNEALLPMLMLLAQAWQSYQSNLSLKDIDIARENDPWILSNKQVGVCVVDLPAQDL 120
Dh 61 NNEALLPMLMLLAQAWQSYQSNLSLKDIDIARENDPWILSNKQVGVCVVDLPAQDL 120
Qy 121 KGLKDKIYFQELGLTYLHMLPEKCEBKGDCGYAVSYRDPNALGTIGLREVIATL 180
Dh 121 KGLKDKIYFQELGLTYLHMLPEKCEBKGDCGYAVSYRDPNALGTIGLREVIATL 180
Qy 181 HEAGISAVVDPIFNHTSHEHMAQRCAGDPLFDNFYTIIPDRRMPDOYDRILREIFPDQ 240
Dh 181 HEAGISAVVDPIFNHTSHEHMAQRCAGDPLFDNFYTIIPDRRMPDOYDRILREIFPDQ 240
Qy 241 HPGGFSQLEDGRWVWTTTNSFCQMDIYNSNPVFWFAMAGEMFLANLGVDIILMDAVAFIM 300
Dh 241 HPGGFSQLEDGRWVWTTTNSFCQMDIYNSNPVFWFAMAGEMFLANLGVDIILMDAVAFIM 300
Qy 301 KQMGTCENLPQAHALLFAFNAVMRIAPAVFPKSEALVHPQVQYIGQDECCQIGYNPL 360
Dh 301 KQMGTCENLPQAHALLFAFNAVMRIAPAVFPKSEALVHPQVQYIGQDECCQIGYNPL 360
Qy 361 QMALIMNTLAVREYVLLHQALTYRHNLPEHTAMVYVSHDDIGWTFADBAAYIGISGY 420
Dh 361 QMALIMNTLAVREYVLLHQALTYRHNLPEHTAMVYVSHDDIGWTFADBAAYIGISGY 420
Qy 421 DHRQFLNRFVNRFDGSPFARGVFPQYNSTGDCRVSTAAALVGLAQDDPHAVDRKILLY 480
Dh 421 DHRQFLNRFVNRFDGSPFARGVFPQYNSTGDCRVSTAAALVGLAQDDPHAVDRKILLY 480
Qy 481 STALSTGGLPLIYLDEVTGLNDNDWSDQNSKSDSRMAHRRRYNEALYAQNDSTAG 540
Dh 481 STALSTGGLPLIYLDEVTGLNDNDWSDQNSKSDSRMAHRRRYNEALYAQNDSTAG 540
Qy 541 QIYQGLRHMLAVROSNPRFDGRLVTFNTNNKHIIIGYIRNNALLAFGNFSEYPTVTAAH 600
Dh 541 QIYQGLRHMLAVROSNPRFDGRLVTFNTNNKHIIIGYIRNNALLAFGNFSEYPTVTAAH 600
Qy 601 LQAMPFKADLIGKTVSLNODLTLOPYQVWMLA 636
Dh 601 LQAMPFKADLIGKTVSLNODLTLOPYQVWMLA 636

```

RESULT 4

```

AA870883
ID AA870883 standard; protein; 856 AA.
XX
AC AA870883;
XX

```

DT 12-JUL-2001 (first entry)

DE Expression vector pSEX-4T-1-AmsUs containing amylosucrase.
 KW Amylosucrase; EC 2.4.1.4; fusion protein; GST; glutathione-S-transferase;
 poly(1,4-alpha-glucan); film production; food additive; cyclodextrin.

OS Unidentified.
 XX WO200125449-A2.
 PN 12-APR-2001.
 PD 04-OCT-2000; 2000WO-EP009695.
 PF 07-OCT-1999; 99DE-01048408.
 PR 07-OCT-1999; 99DE-01048408.
 XX

PA (AXIVA-) AXIVA GMBH.

PI Bengs H, Polakowski T, Held A, Gallert K;
 XX WPI; 2001-328330/34.

DR N-PSDB; AAF61712.

PT Amyloisurase immobilized as fusion protein with anchoring sequence,
 PT useful in continuous preparation of poly(1,4-alpha-glucan) from sucrose.
 XX
 PS Claim 2; Page 34-38; 38pp; German.

This invention describes a novel amylosucrase (AS), immobilizable on a solid phase, which comprises a fusion protein (FP) of functional units of AS, an anchoring sequence, and optionally additional auxiliary sequences. The invention also describes (1) nucleic acid (I) encoding FP; (2) expression vector containing (I) and able to express FP in a host cell; *Escherichia coli* containing the vector of (2); (3) anchoring sequence, or its functional variants or fragments, of at least 8 nucleotides that encodes an epitope, a high-affinity binding partner or GSP (glutathione-S-transferase); (4) solid phase for immobilizing AS comprising glutathione-S-sepharose; (5) combined, stable catalyst (A) comprising FP immobilized on Sepharose for production of poly(1,4-alpha-glucan) (II); and (6) biocatalytic production of (II). AS is used in production of poly(1,4-alpha-glucans) (II), useful for producing films, as food additives, as starting materials for cyclodextrins and as auxiliaries in pharmaceutical formulations. Immobilized AS makes possible efficient, inexpensive and continuous production of poly(1,4-alpha-glucan) (II), and it can be used repeatedly. Compared with known methods, specifically is improved (increased yield of (II) and reduced formation of palatinsose) and reaction is complete within 24 hours, compared to 48-72 hours for batch methods. This sequence represents the expression vector construct pGEX-4T1-1.Amsus which contains the amylosucrase (BC 2.4.1.4) described in the invention

SQ Sequence 856 AA;

Query Match 98.6%; Score 3360; DB 4; Length 856;

Best Local Similarity 99.8%; Pred. No. 0;

Matches 626; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy	10	LIILOYKTRLIDITYEBQAGIEKSEDMQSRBMDHPKPLXMEBLSVYGNNAIILPML	69
Oy	70	EMLLAQMOSSYSGRNSLSKIDIDIAENNDWITLSKQGVGYCYDLFPAGDKJLKDPIY	129
Dp	230	MIIQYKTRLLDITYEBQAGIEKSEDMQSRBMDHPKPLXMEBLSVYGNNAIILPML	289
Oy	290	EMLLAQMOSSYSGRNSLSKIDIDIAENNDWITLSKQGVGYCYDLFPAGDKJLKDPIY	349
Dp	130	FOELGLTYILMPLFKCPEKSGSDGVAVSSYADVNPALGTIGDLREVIAALHEAGISAVY	189
Oy	350	FOELGLTYILMPLFKCPEKSGSDGVAVSSYADVNPALGTIGDLREVIAALHEAGISAVY	409
Dp	190	DFIFNHTSNHEMAQCAAGDPLFDNFYIIPDRMPQYRTLRRLTLPDQHPCGSQLE	249
Oy	410	DFIFNHTSNHEMAQCAAGDPLFDNFYIIPDRMPQYRTLRRLTLPDQHPCGSQLE	469
Dp	250	DGRNWTTESSPQMDLTVNSNPWFYFAMGEMFLANLGVDLIRMDAVAFIMKQGSTCEN	309
Oy	470	DGRNWTTESSPQMDLTVNSNPWFYFAMGEMFLANLGVDLIRMDAVAFIMKQGSTCEN	529
Dp	310	LPQHALIRAFNAVMEIAAPAVEFKSEAIVHPDQVOYITGODECOIGINPLQAMLLMNTL	369
Oy	530	LPQHALIRAFNAVMEIAAPAVEFKSEAIVHPDQVOYITGODECOIGINPLQAMLLMNTL	589
Dp	370	ATREVNLLHQALTYRRNLPEHTAMVYVYVSHDICTGTFADBDAAVLISGYDRHQLNRF	429
Oy	590	ATREVNLLHQALTYRRNLPEHTAMVYVYVSHDICTGTFADBDAAVLISGYDRHQLNRF	649
Dp	430	FVNRFDGSPKRGVPEQINPSTGDCRVSGTAAALVGLAODDPRADVRLTILYSIALSTGGL	489
Oy	650	FVNRFDGSPKRGVPEQINPSTGDCRVSGTAAALVGLAODDPRADVRLTILYSIALSTGGL	709
Dp	490	PLIYLGEBEVETLNDDDMSODSNKSDSRBAHPRYNBAIYVQORNDPSTPAQOIYQGLRHM	549
Oy	710	PLIYLGEBEVETLNDDDMSODSNKSDSRBAHPRYNBAIYVQORNDPSTPAQOIYQGLRHM	769
Dp	550	IAYRQSNPREDEGRGLTVETNNKHIIGYTRNNALAFQNFSEYDPYVVAHTLOAMPKAH	609

Db	Qy
770	610
I A V R S N P R F D G G R L V T F E N I N K H K I I G Y I R N N A L L A F G F S E Y P Q V T A H T L Q A M P E K A H	D L I G K T V S I N D L T I Q P Y Q V W M M L E I A
829	636
	830
	D L I G K T V S I N D L T I Q P Y Q V W M M L E I A
	856

RESULT 5

AAR88386

ID AAR88386 standard; protein; 614 AA

1
 2
 3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525

AC AAR88386;

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

DI 15-OCT-1996 (first entry
YY

Not acceptable for research and

LE
XX
NELSBELLA POLYBACCINATED A

amy]ovisrags: enzyme: amy]
KW

EC-2, 4-1, 4: crop improvement

XXIX

Neisseria polysaccharea.

XX

PN W09531553-A1.

XX

PD 23-NOV-1995.

XX

PF 18-MAY-1995; 95WO-EP001

XXXXXX

PR 18-MAY-1994; 94DE-04417

PR 22-DEC-1994; 94DE-044447

XX
XX
(CIVIL) TYPE CERTIFICATEPA (GENB-) INST GENBIOLOGISC
XX

XX	Kocswann J	Buettcher W
PI		

BUCHNER V,
ROSSMANN U,
PL
XX

WPB: 1996-010938/01

DR NEI, 1998-010330/VI;
N-PSDB: AAT08960 -

[illegible]

DNA encoding amylo:sucrase

linear 1,4-glucan(s), esp

11-11-11

PS Claim 1; Page 39-42; 56pp

Query Match	Best Local Similarity	74.5%;	Score 2540;	DB 2;	Length 614;	
Matches	494;	Conservative	13;	Mismatches 43;	Indels 52;	Gaps 6
QY	1	MLPTQVGLILQYIKRILDIYYPEQAGIEKSEDMQFOSRRMDTHFPKLMNELDSYVG	60			
DB	1	MLPTQVGLILQYIKRILDIYYPEQAGIEKSEDMQFOSRRMDTHFPKLMNELDSYVG	60			
QY	61	NNELALLPELLEMLLQAWOMYSQORNSLSKLDIDIAEENNDWLLSNKQVAGVCYVDLFAGDL	120			
DB	61	NNELALLPELLEMLLQAWOMYSQORNSLSKLDIDIAEENNDWLLSNKQVAGVCYVDLFAGDL	120			
QY	121	KGLDKDIPIYFQELGLTYVDLMLPLFKCEPGEKSGDGYAVSSSYRDVNPALGTIGDLREVIAAL	180			
DB	121	KGLDKDIPIYFQELGLTYVDLMLPLFKCEPGEKSGDGYAVSSSYRDVNPALGTIGDLREVIAAL	180			
QY	181	HEAGISAVVDIPIFMTSNEHEMAQRCAAGDLFNFXYIIFPDRMPDQYDRTLREIIPDQ	240			
DB	181	HEHSHRRRRFLYPLQLQRT-RMAQR-C-AGDPLFNFXYIIFPDRMPDQYDRTLREIIPDQ	238			

```

OY 241 HPGGFQLEDGRWVWTTTNSFQMDLNSNPWF---RAMAGMLFLANLGVDLRMADA 297
DB 239 HPGGFQLEDGRWVWTTTNSFQMDLNSNPWFPAQMRKCCSLPTWALTSCVWMLPLRG 298
OY 238 FIMKMGTSCEMLPOAHALIRAFNAVMRIADAVFEKSEALVHPDQVVOYIGODECQIGY 357
DB 229 NKMGGAAKTC---AAHALIRAFNAVMRIADAVFEKSEALVHPDQVVOYIGODECQIGY 354
OY 358 NPLQMALIMNTLATREVMILHQAALTYRNLPHETMVMVVRSHDIDGTFADADAAYLGI 417
DB 355 NPLQMALIMNTLATREVMILHQAALTYRNLPHETMVMVVRSHDIDGTFADADAAYLGI 414
OY 418 SGYDRQFLNRFVNRPDGSPFARGVFOYNPSTGDCRVSTAAALVGLAQDDPHAVDRIX 477
DB 415 SGYDRQFLNRFVNRPDGSPFARGVFOYNPSTGDCRVSTAAALVGLAQDDPHAVDRIX 474
OY 478 LLYSTALSTGGLPLIYLGDVEGTLNDDWSQDSNKSDDSRMAHRRRYNEALYAQRNDPST 537
DB 475 LLYSTALSTGGLPLIYLGDVEGTLNDDWSQDSNKSDDSRMAHRRRYNEALYAQRNDPST 512
OY 538 AAG-----QIYQGLRMIAVRQSNPRFPDGRGLVTFNTNNKHITG 576
DB 513 AVGPSPALQRPVATERSVDNRQIYQGLRMIAVRQSNPRFPDGRGLVTFNTNNKHITG 572
OY 577 YI 578
DB 573 YI 574

RESULT 6
AAR8633
ID AAR8633 standard; protein; 614 AA.
AC AAR8633;
XX 20-JUN-1996 (first entry)
XX Neisseria polysacchara amylosucrase.
DE Neisseria polysacchara amylosucrase.
XX Amylosucrase; bacteria; fungi; plants; detection; transformation; linear;
XX alpha-1,4-glucans; amylose; sucrose; colourless; odourless; tasteless;
XX non-toxic; biodegradable; self-sustaining; films; fibres; textiles;
XX paper-making; glass-fibre; tablet binder; food thickener; sound proofing;
XX flow properties; paraffin oils; organic compound inclusion;
XX chromatographic separation; cyclodextrins.
XX
XX Neisseria polysacchara.
XX
XX DE4417879-A1.
XX
XX 23-NOV-1995.
XX
XX 18-MAY-1994; 94DE-04417879.
XX
XX 18-MAY-1994; 94DE-04417879.
XX
XX (GENB-) INST GENBIOLOGISCHE FORSCHUNG.
XX
XX Kossman U, Buettcher V, Welsh T;
XX
XX MPI; 1996-000447/01.
XX
XX N-PSDB; AAT11179.
XX
XX New DNA sequence encoding amylo:sucrase of Neisseria - and transformed
XX plant, bacteria and fungi able to produce linear alpha-1,4-glucan(s),
XX esp. amylose, in practically pure form.
XX
XX Claim 5; Page 23-27; 42pp; German.
XX
XX The N. polysacchara DNA sequence AAT11179, which encodes AAR8633
XX amylosucrase (ASA), can be used to produce bacteria, fungi and plants
XX that express ASA, and to detect and isolate related DNA from other

```

```

CC organisms. Transformed plants which express ASA are able to produce
CC linear alpha-1,4-glucans, specifically amylose from sucrose, which can be
CC used to produce colourless, odourless, tasteless, nontoxic,
CC biodegradable, self-sustaining films or fibres, e.g. for use in the food,
CC textile, paper-making and glass-fibre industries. Amylose can also be
CC used as a binder for tablets, thickener for food, in sound proofing
CC panels, to improve flow properties in paraffin-based oils, for inclusion
CC of organic cpds., in chromatographic sepn. and as a starting material for
CC cyclodextrins
XX
SQ Sequence 614 AA;
Query Match 74.5%; Score 2540; DB 2; Length 614;
Beet Local Similarity 82.1%; Pred. No. 5,1e-239;
Matches 494; Conservative 13; Mismatches 43; Indels 52; Gaps 6;
OY 1 MLTPYQVGLILQYKTRILDIYPEORAGLEKSEDMQFSRRMTHPKLNNELDSVYG 60
DB 1 MLTPYQVGLILQYKTRILDIYPEORAGLEKSEDMQFSRRMTHPKLNNELDSVYG 60
OY 61 NNEALLPMLIEMILQAQMSYSGRNSLKDIDIARENNPDWILSNQVGVVYVDLFAQDL 120
DB 61 NNEALLPMLIEMILQAQMSYSGRNSLKDIDIARENNPDWILSNQVGVVYVDLFAQDL 120
OY 121 KGLKDKIPYQELGLTYLHMLPFLKCPKSGSDGYAVSSYRDVNPALGTTGLREYIAL 180
DB 121 KGLKDKIPYQELGLTYLHMLPFLKCPKSGSDGYAVSSYRDVNPALGTTGLREYIAL 180
OY 181 HBAQSAVVDPIFNNTSNHBAQRCAGDPLFDNFYIYFPRRMPDQDRLREIFPQ 240
DB 181 HESHFRRRFYIQPLQRT-RMAQRC-AGDPLFDNFYIYFPRRMPDQDRLREIFPQ 238
OY 241 HPGGFQLEDGRWVWTTTNSFQMDLNSNPWF---RAMAGMLFLANLGVDLRMADA 297
DB 239 HPGGFQLEDGRWVWTTTNSFQMDLNSNPWFPAQMRKCCSLPTWALTSCVWMLPLRG 298
OY 238 FIMKMGTSCEMLPOAHALIRAFNAVMRIADAVFEKSEALVHPDQVVOYIGODECQIGY 357
DB 229 NKMGGAAKTC---AAHALIRAFNAVMRIADAVFEKSEALVHPDQVVOYIGODECQIGY 354
OY 478 LLYSTALSTGGLPLIYLGDVEGTLNDDWSQDSNKSDDSRMAHRRRYNEALYAQRNDPST 537
DB 475 LLYSTALSTGGLPLIYLGDVEGTLNDDWSQDSNKSDDSRMAHRRRYNEALYAQRNDPST 512
OY 538 AAG-----QIYQGLRMIAVRQSNPRFPDGRGLVTFNTNNKHITG 576
DB 513 AVGPSPALQRPVATERSVDNRQIYQGLRMIAVRQSNPRFPDGRGLVTFNTNNKHITG 572
OY 577 YI 578
DB 573 YI 574

RESULT 7
AAU5090
ID AAU5090 standard; protein; 640 AA.
AC AAU5090;
XX 27-FEB-2002 (first entry)
XX
XX Propionibacterium acnes immunogenic protein #10986.
XX
XX SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
XX uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;

```

KM	inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
KM	dermatological; osteopathic; neuroprotectant.
XX	
OS	Propionibacterium acnes.
XX	
EN	WO200181581-A2.
XX	
PD	01-NOV-2001.
XX	
PF	20-APR-2001; 2001WO-US012865.
XX	
PR	21-APR-2000; 2000US-0199047P.
XX	
PR	02-JUN-2000; 2000US-0208841P.
XX	
PR	07-JUL-2000; 2000US-0216747P.
XX	
PA	(CORI-) CORIXA CORP.
PI	Skeiky YAM, Persing DH, Mitcham JL, Wang SS, Bhattacharya A,
PI	L'maisonneuve J, Zhang Y, Jen S, Carter D;
XX	
DR	WPI: 2001-616774/71.
XX	N-PSDB; AAS59546.
PT	Propionibacterium acnes polypeptides and nucleic acids useful for
PT	vaccinating against and diagnosing infections, especially useful for
PT	treating acne vulgaris.
PS	
XX	
PS	Example 1; SEQ ID NO 11285; 1069pp; English.
XX	
CC	Sequences AAU39105-AAU6017 represent Propionibacterium acnes immunogenic
CC	polypeptides. The proteins and their associated DNA sequences are used in
CC	the treatment, prevention and diagnosis of medical conditions caused by
CC	P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
CC	pustulosis, hyperostosis and osteomyelitis), uveitis and endophthalmitis.
CC	P. acnes is also involved in infections of bone, joints and the central
CC	nervous system, however it is particularly involved in the inflammatory
CC	lesions associated with acne vulgaris. A method for detecting the
CC	presence or absence of P. acnes in a patient comprises contacting a
CC	sample with a binding agent that binds to the proteins of the invention
CC	and determining the amount of bound protein in the sample. The
CC	polypeptides may be used as antigens in the production of antibodies
CC	specific for P. acnes proteins. These antibodies can be used to
CC	downregulate expression and activity of P. acnes polypeptides and
CC	therefore treat P. acnes infections. The antibodies may also be used as
CC	diagnostic agents for determining P. acnes presence, for example, by
CC	enzyme linked immunosorbent assay (ELISA). Note: The sequence data for
CC	this patent did not form part of the printed specification, but was
CC	obtained in electronic format directly from WIPO at
CC	ftp.wipo.int/pub/published_pct_sequences
XX	
XX	Sequence 640 AA;
Query Match	34.2%; Score 1164; DB 4; Length 640;
Best Local Similarity	39.4%; Pred. No. 2.9e-104;
Matches 245; Conservative 113; Mismatches 230; Indels 34; Gaps 12	
QY	38 RQFSRMDTHPEKKNLELDVYGN--NEALFLMLELTLAQAWQSYORSRSLKIDIDARE 95
DB	21 QSFDRMWEHYLDLITGSRVYGDRADELTLQIRAILLRKC---AERDIDLRKIDEARL 76
QY	96 NNPDWLISNKKQYGVGYCVYDLFAGDRLKGLKDKPIPYQELGLTYLHMLPKCEBGSQGY 155
DB	77 LEPDWLQOPWMLGVMTYDTHFSGTLKGISDHDHLCMDGVRYLHMLPILQOPGQTDGXY 136
QY	156 AVSSRTDVNPPALGTIGDLREVLTAALHEAGISAVVDFIFPHITSNHHMAORCAAGDLEFN 215
DB	137 AVADERTIRTDGTMDLADLTATIRAHGISLVMDLIVNVAHEHMAARARAGQOKYND 196
QY	216 FYIYIPDDRMPOCYORTLEIRIPDQHPGFSOLEDR--WVWTFPNSFOQDLNYSNPWYR 274
DB	197 YPHILSTQDEVDAEMKLPDVFPPRAHNGFTYDDDDCGGVMAATPNEFGQDLWMAANDVFC 256
QY	275 AMAGMFLFANIGVDILRMDAVAFIMKQNGTSCENLPQAHILIRAFNAVRITAAPAVEFK 334

[illegible]


```

QY 290 IIRMDVAATIKWQMTSCENTLPOAHLIRAFNAWRIAPAYFEKSEATVHPDQVYQITG 349
DB 95 VRLDIAIATIKKLTGTCNQLPEIHDTQSLRQAIRAPAVAFADALVGDDDLTYGFG 154
QY 350 QDE-----COIGNPLQMALMNTLATREVNLIHQALTFRNHLPEHTAVNYYVRSDDIG 404
DB 155 RGRHMKVCYDMTHNSLMVQMSALATRVSLMERTLSLSTPPKSTTTTATYAKCHDDIG 214
QY 405 WTPADEDAAYLIGSYDHRQPLNRFVNRFDGSFARGVDFQYNPSTGDCRSVSTAALVG 464
DB 215 WTVVDADARKTGIDPFAHQPLSDFFVSGTFPPSPFANGVDFQNPPTGDRIRIGSLASLAG 274
QY 465 ----LAODDPHAVD-----RIKLLYSIALSTGCLPLIYLDVGTGLNDDWSQDSNKSDDS 516
DB 275 LBSALSDPDPAGVDAIARIIVMLHTALIGYGVPLIMGDEYGMGLN-DWMORDPGHADPN 333
QY 517 RAAHPRRYEALYAQGN-DPSTAAGQIYQGLRHMTLAVRQSNRFDGS-RLVTFNTNNKH 574
DB 334 RWHHRPMMNWSVWQKAAHAPHSPGATIMGVRRAINARHSPEFHASVDTTVVLSEPHKV 393
QY 575 IGYIR---NNALLAFGNFSEYPOVTYTAHTLQAMPFXADLLIGKTVSLNODLT-----LQ 626
DB 394 IWMGRPHPEGRMIELYNISEHEWVFEMETLRS---ELDDVYVELLRGFDYDLTPMMLRLA 450
QY 627 PYQVWML 633
DB 451 PYECIML 457

RESULT 10
ID AAU60953 standard; protein; 622 AA.
AC AAU60953;
XX
DT 27-FEB-2002 (first entry)
XX
DB Propionibacterium acnes immunogenic protein #21849.
XX
KM SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
KM uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
KM inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
KM dermatological; osteopathic; neuroprotectant.
XX
OS Propionibacterium acnes.
XX
PN WO200181581-A2.
XX
PD 01-NOV-2001.
XX
PF 20-APR-2001; 2001WO-US012865.
XX
PR 21-APR-2000; 2000US-0199047P.
PR 02-JUN-2000; 2000US-0208841P.
PR 07-JUL-2000; 2000US-0216747P.
XX
PA (CORI-) CORIXA CORP.
XX
PI Skelky YAM, Persing DH, Mitcham JL, Wang SS, Bhatia A;
PI L'maisonneuve J, Zhang Y, Jen S, Carter D;
XX
DR WPI, 2001-616774/71.
DR N-PSDB; AAS59613.
XX
PT Propionibacterium acnes polypeptides and nucleic acids useful for
PT vaccinating against and diagnosing infections, especially useful for
PT treating acne vulgaris.
XX
PS Example 1; SEQ ID NO 22148; 1069bp; English.
XX
CC Sequences AAU9105-AAU68017 represent Propionibacterium acnes immunogenic
CC polypeptides. The proteins and their associated DNA sequences are used in

```

```

CC the treatment, prevention and diagnosis of medical conditions caused by
CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
CC pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.
CC P. acnes is also involved in infections of bone, joints and the central
CC nervous system, however it is particularly involved in the inflammatory
CC lesions associated with acne vulgaris. A method for detecting the
CC presence or absence of P. acnes in a patient comprises contacting a
CC sample with a binding agent that binds to the proteins of the invention
CC and determining the amount of bound protein in the sample. The
CC polypeptides may be used as antigens in the production of antibodies
CC specific for P. acnes proteins. These antibodies can be used to
CC downregulate expression and activity of P. acnes polypeptides and
CC therefore treat P. acnes infections. The antibodies may also be used as
CC diagnostic agents for determining P. acnes presence, for example, by
CC enzyme linked immunosorbent assay (ELISA). Note: The sequence data for
CC this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 622 AA;

```

```

Query Match 16.4%; Score 560.5; DB 4; Length 622;
Best Local Similarity 28.6%; Pred. No. 3.5e-45;
Matches 162; Conservative 85; Mismatches 220; Indels 99; Gaps 22;

```

```

QY 118 GDLKGLKRTIPYFGLGLTYHMLPFCRPGKSGGSAVSSYRVNPAITGDLREVI 177
DB 104 GDFKGLTGKLDYLOWLGVDCMLPFPYDSP-LHGGYDIRRYRWIRELGTIEFQVFL 161
QY 178 AALHEAGISAVVDPIFNTHSNHEWQRCAG-DELFNFFYIFPDRRMPDQYDLREI 236
DB 162 DAHNRGRRLVILDFPMNHTSDSHFPQSSRADPDQPYGN-YVWMSD--TDAAYDARI 217
QY 237 PEDQHPGFSQLEDGRWVW-TTFNSFQW-----DINTSNPVWFRAMAGMLFLANIG 287
DB 218 FCD-----TEDSNMGWDSQRQFYWHRRFHHQPDINFEPRVWEEMLDVRFWMDLG 269
QY 288 VDILMDVAATIKWQMTSCENTLPOAHLIRAFNAWRIAPAYFEKSEATVHPDQVYQ 347
DB 270 IDGFLDVAVPYLIEAGNCCENTLPGTHKILQRLAMVBEYFGRILLCEANQMPDQV 329
QY 348 IGO--DECOIGNPLQMALMNTLATREVNLIHQALTFRNHLPEHTAVNYYVRSDDIGWT 406
DB 330 FONGDECOMAHHFPVMPRLYMGRLGRSRECSILAAHPHLPDGCQMGFTLRNHELTJE 389
QY 407 F-ADEDAAYLIGSYDHRQPLNRFVNRFDGSFARGVDFQYNP-STGDCRSVSTAALVG 464
DB 390 MYTEEDRHVW-----EYAFESRMRCNL-GIRRRRLSP 421
QY 465 LAODDPHAVDRIKLLYSIALSTGCLPLIYLDVGTGLNDDWSQDSN-----KSDDSRW 518
DB 422 LVNDND--AKIRILNAMLSTLPGSPVLYYDDEIG-MQDDPWLPRDGVRTFMQDDSET 477
QY 519 A-----HRPRY-NEALYAQRNDPSTAAGQIYQGLRHMTLAVRQSNR 558
DB 478 AGFSTALPEDRHLPRIKTFGHDPEHVNVA--RQMDPDS-----LLVWRAMLGIRRHV 531
QY 559 FDGGLVTFNTNNKHIIIGYIRNN--ALLAFGNFSEYPOVTYTAHTLQAMPFXADLLIGSK 615
DB 532 FGTGFTDLGGEDMAVMSFLRHNEHRTVLCIANFSDTERMALHLPORAGMTGSSLIHQ 591
QY 616 TVS-LNODLTQ----PYQVWML 636
DB 592 DAQPVKADGTLSVPLMPYGRWLOVS 617

RESULT 11
ID ABM57472 standard; protein; 622 AA.
AC ABM57472;
XX ABM57472;
XX
DT 20-OCT-2003 (first entry)

```

XX DE Propionibacterium acnes predicted ORF-encoded polypeptide #22148.
XX KW Acne vulgaris; antibeorthoeic; dermatological; antibacterial;
XX KM immunostimulant; immune response; vaccine.
XX OS Propionibacterium acnes.
XX PN WO2003033515-A1.
XX PD 24-APR-2003.
XX PF 11-OCT-2002; 2002WO-US032727.
XX PR 15-OCT-2001; 2001US-00978825.
XX PA (CORI-) CORIXA CORP.
XX PI Mitcham JY, Skeiky YAM, Persing DH, Bhatia A, Maisonneuve JY;
PI Zhang Y, Wang S, Jey S, Lodes MJ, Benson DR, Jones R, Carter D;
PI Barth B, Valliave-Douglas J;
XX DR WPI; 2003-381789/36.
XX DR N-PSDB; ACF64542.
XX PT New Propionibacterium acnes polypeptides and polynucleotides encoding the
XX PT polypeptide, useful for diagnosing, preventing or treating acne vulgaris,
XX PT or for stimulating an immune response specific for a P. acnes protein.
XX PS Example 1; SEQ ID NO 22148; 1481bp; English.
XX CC The invention relates to an isolated polynucleotide (ACF64435-ACF64733)
XX CC encoding a Propionibacterium acnes protein. The invention also relates to
XX CC polypeptides encoded by the polynucleotides (ABM35624-ABM64536) and to
XX CC immunogenic fragments of P. acnes polypeptides. The invention
XX CC additionally encompasses expression vectors and host cells comprising a
XX CC polynucleotide of the invention; antibodies against polypeptides of the
XX CC invention; fusion proteins comprising a polypeptide of the invention; a
XX CC method for stimulating an immune response specific for a P. acnes
XX CC polypeptide and an isolated T cell population comprising T cells prepared
XX CC via this method; a vaccine composition (comprising P. acnes polypeptides,
XX CC polynucleotides, antibodies, fusion proteins, T cell populations, or
XX CC antigen-presenting cells that express the polypeptide); a method and kit
XX CC for detecting or determining the presence or absence of P. acnes in a
XX CC patient; and a method for inhibiting the development of P. acnes in a
XX CC patient. The P. acnes polypeptides, polynucleotides, antibodies, fusion
XX CC proteins, T cell populations or antigen-presenting cells that express the
XX CC polypeptides are useful for diagnosing, preventing or treating acne
XX CC vulgaris, or for stimulating an immune response specific for a P. acnes
XX CC protein. The polynucleotides can also be used as probes or primers for
XX CC nucleic acid hybridisation. The vaccine composition is useful for the
XX CC stimulation of an immune response against P. acnes, or for treating acne,
XX CC and the kit is useful for performing a diagnostic assay. The present
XX CC sequence represents a polypeptide predicted to be encoded by an ORF (open
XX CC reading frame) contained within the P. acnes polynucleotides of the
XX CC invention. Note: The sequence data for this patent did not form part of
XX CC the printed specification, but was obtained in electronic format directly
XX CC from WIPO at ftp.wipo.int/pub/published_pcc_sequences
XX SQ Sequence 622 AA;
Query Match 16.4%; Score 560.5; DB 6; Length 622;
Best Local Similarity 28.6%; Pred. No. 3.5e-45;
Matches 162; Conservative 85; Mismatches 220; Indels 99; Gaps 22;
QY 118 GDLKGIKDKIPYFOELGTLYIHLMLPLFKCPEKSGDGYAVSYRDVNPALGTIGLREVI 177
DB 104 GDFXGUTGDLADYLQMLGVDCMLWLPFYDGP--LHOGYDIRDYRWLRRELTGTEPKVPL 161
QY 178 AALHAGISAVVDIFFNHTSNHEWAGRCAG--DPLFDNFYITFDRRMPDQYDRTIKRI 236
DB 162 DAHHRGRLVILIDFVNMHTSDSHWFOSSRADPDGPGYGN--YVWSD--TDRAVSDARTI 217

QY 237 FPDQHPGGSFQLEDGRWYV--TFNSFQW-----DUNSNPWYFRAMAGMLFLANLG 287
DB 218 FCD-----TEDSNWSHDSQKQFYHFRFHQPDLANFEPRVMEKLDVRFMDUG 269
QY 288 VDLRMDAVAFIWKQMGTSCEMLPOAHLLIRAFNAVMRIAABVFEKSEAIYHPQVYQY 347
DB 270 IDGFRLDANVPYIIEAGTNCENTLPGTHKLKQLRAMVBEYGRILLCEANQMPDQVVEY 329
QY 348 IQG-DECGIYNPLQMLAMTLATREVNLLHQALTTRNLPEHRAVMYVYRSHDIDGT 406
DB 330 FGNDECCQWAFHFVYMPRLYMGIRGSRCEISILLAPHPIDGCGWGTFLRNHDLTLE 389
QY 407 F-ADEDAVYAGISGYDRHQFLARFFVNRPDGSFARGVPRQYNP--STGDCRVGTAALVY 464
DB 390 MYTEEDRHRYM-----EETAPESKRCNL--GIRRRISP 421
QY 465 IAAODPHAVDRIKLILYSIALSTGGLPLIYAGDEVGLTANDDMSODS-----KSDDSRW 518
DB 422 LVDND--AKIRLNLAMLLSLGSPVLYYDEIG--MGDDPMLPDRDGVRTPMQMDSDT 477
QY 519 A-----HPRY--NEALVYQRNDPSTAAQIYQGLRHMTAVRQSNR 558
DB 478 AGFSTALPEDFHLPLIRTFGHDPEHVNA--ROMDDPS--LWVTRAMLGIRRHBY 531
QY 559 FDGRLVTFNTNKHILIGYIRNN--ALLAFGNFSEYPTVTAHTLOAMPFAHDLIGX 615
DB 532 FTGGEFTDLCGPMAVMSFLRNEHETVCLANFSTERMAHLHPQFAMTGSSLIHQ 591
QY 616 TVS-IMQDLTLO---PYQVMLEIA 636
DB 592 DAQPVKADGTLVPLMPYGYRWLGVS 617
RESULT 12
ABU39607
ID ABU39607 standard; protein; 1106 AA.
XX AC ABU39607;
XX DT 19-JUN-2003 (first entry)
XX DE Protein encoded by Prokaryotic essential gene #25134.
XX KW Antisense; prokaryotic essential gene; cell proliferation; drug design.
XX OS Pseudomonas putida.
XX PN WO200271183-A2.
XX PD 03-OCT-2002.
XX PF 21-MAR-2002; 2002WO-US009107.
XX PR 21-MAR-2001; 2001US-00815242.
XX PR 06-SEP-2001; 2001US-00948993.
XX PR 25-OCT-2001; 2001US-0342923P.
XX PR 08-FEB-2002; 2002US-00072851.
XX PR 06-MAR-2002; 2002US-0362699P.
XX PA (ELIT-) ELITRA PHARM INC.
XX PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyekind JW;
PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX DR WPI; 2003-029926/02.
XX DR N-PSDB; ACA43477.
XX PT New antisense nucleic acids, useful for identifying proteins or screening
XX PT for homologous nucleic acids required for cellular proliferation to
XX PT isolate candidate molecules for rational drug discovery programs.
XX XX Claim 25; SEQ ID NO 67531; 1766bp; English.

The invention relates to an isolated nucleic acid comprising any one of the 6213 antisense sequences given in the specification where expression of the nucleic acid inhibits proliferation of a cell. Also included are: (1) a vector comprising a promoter operably linked to the nucleic acid encoding a polypeptide whose expression is inhibited by the antisense nucleic acid; (2) a host cell containing the vector; (3) an isolated polypeptide or its fragment whose expression is inhibited by the antisense nucleic acid; (4) an antibody capable of specifically binding the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular proliferation or the activity of a gene in an operon required for proliferation; (7) identifying a compound that influences the activity of the gene product or that has an activity against a biological pathway required for proliferation, or that inhibits cellular proliferation; (8) identifying a gene required for cellular proliferation or the biological pathway in which a proliferation-required gene or its gene product lies or a gene on which the test compound that inhibits proliferation of an organism acts; (9) manufacturing an antibiotic; (10) profiling a compound's activity; (11) a culture comprising strains in which the gene product is overexpressed or underexpressed; (12) determining the extent to which each of the strains is present in a culture or collection of strains; or (13) identifying the target of a compound that inhibits the proliferation of an organism. The antisense nucleic acids are useful for identifying proteins or screening for homologous nucleic acids required for cellular proliferation to isolate candidate molecules for rational drug discovery programs, or for screening homologous nucleic acids required for proliferation in cells other than *S. aureus*, *S. typhimurium*, *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of the target prokaryotic essential genes. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences

Sequence 1106 AA:

Query Match 16.1%; Score 550; DB 6; Length 1106;
Best Local Similarity 29.0%; Pred. No. 9.5e-44;
Matches 164; Conservative 83; Mismatches 228; Indels 90; Gaps 20;

118 GDUKGLKDKIPYQEGLYTHLMPLKCEKSGSGYAVSSRDVNPALGTIGDREYI 177
39 GPGAGHISKLDYTAELGVMTLWLPY--PSPRRDGYDIAEYKAVHPYGSAAARRPI 96
178 AALHEAGISAVVDFINHTSNEHEMAOR---CAAGDPLFNFYIIPDDRMPPQ--YDRIL 233
97 AEAHKGRLVYTELVINHTSDQHPFORAHAKGSKARE--FYVSD--DDQKXGDI- 150
234 REIFPDHPGGS-QLEDGRVWTTTNSFQMDLNSNFWFRAMAGEMFLANLGVDTIR 292
151 RIIFLDEKSNMTWDPVAGGYFMRHRYSHQPDLPNPOYLKAVIGMFPMLDLGVGJR 210
293 MDAVATWQMGTSCEMLPQAAALIRAFNAVRIAPVFFKSEALVHPDQVYVIGQ-- 350
211 LDAIPYLIERDGTNNELAEHDLVLAIRAEIDANYPDRMLAEANQWPEDETPRYGEGD 270
351 -DECOJGVNPLQALIMNTLATREVNLMHQAITYRNHLPEHTAWVNVVSHDIDGWTFF-A 408
271 GDECHMAHFPPLMPKRMALAMEDRPITDILKQTEIPANCMAFLRNHDELTEMT 330
409 DEDAAVYIGISGVDRHOFNLNRFVNRPDGSPARGVPPOYNPSTDCRVS--GTAALVGLAQ 467
331 DRRERDL-----NNVYAEERRAIIINGIRRLAPLQ 362
468 DDPHAVDRITKLVSIALSTGGLPLIYGDEVGT-----LNDQ-----WSQDSN----- 511
363 RDRR---RIELITSLISMGPFTLIYGGELGMDNIIYLGBRDGVATPQWMSDRNGGGS 419
512 KSPDSMAHRRPRNEALY-----AQRNDPSTAAGIYQGRHMLAVQSPRPDGG 563
420 RADPQVLV-LPPTMDPLGYQTVNVEAQSHDHS---LNTWRRLMAVRKQKQKAGRG 474
564 LVTFNTNKKIIGYIR-----NNALLAFGNFSEVPQTVAHTLIQAMFKAHDIIGCK 615
475 LRTLTSPNRIILAVIREYTDADGNTVEILCVANVSRAQAALDELQGYADKVEVEMIGGS 534

QY 616 TV-----SINQDITLQPYQVWMEIA 636
DB 535 APPPIGQLEPFLTLPEYAFYFLLA 559

RESULT 13

AAU36268
ID AAU36268 standard; protein, 1100 AA.

AAU36268;

14-FEB-2002 (first entry)

Pseudomonas aeruginosa cellular proliferation protein #258.

Antisense; prokaryotic cellular proliferation protein; antibiotic;
antibacterial; drug design.

Pseudomonas aeruginosa.

WO200170955-A2.

27-SEP-2001.

21-MAR-2001; 2001WO-US009180.

21-MAR-2000; 2000US-0191078P.

23-MAY-2000; 2000US-0206848P.

26-MAY-2000; 2000US-0207727P.

23-OCT-2000; 2000US-0242578P.

27-NOV-2000; 2000US-0253625P.

22-DEC-2000; 2000US-0257931P.

16-FEB-2001; 2001US-0269308P.

(ELITR-) ELITRA PHARM INC.

Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GD;
Yamamoto RT, Xu HH;
WPI; 2001-611495/70.

N-PFDB; AAS54127.

New polymucleotides for the identification and development of
antibiotics, comprise sequences of antisense nucleic acids.

Example 3; SEQ ID NO 11861; 511pp; English.

The invention relates to antisense inhibitors of genes essential to
prokaryotic cellular proliferation, their use in identifying the genes,
their use in the discovery of novel antibiotics, the essential genes
themselves and the encoded proteins. The prokaryotes used are *Escherichia*
coli, *Staphylococcus aureus*, *Salmonella typhi*, *Klebsiella pneumoniae*,
Pseudomonas aeruginosa and *Enterococcus faecalis*. The invention is also
useful for the identification of potential new targets for antibiotic
development. The antisense nucleic acids can also be used to identify
proteins used in proliferation, to express these proteins, and to obtain
antibodies capable of binding to the expressed proteins. The proteins can
be used to screen compounds in rational drug discovery programmes. The
antisense nucleic acid sequence is also useful to screen for homologous
nucleic acids which are required for cell proliferation in a wide variety
of organisms. The present sequence represents an essential prokaryotic
cellular proliferation protein. Note: The sequence data for this patent
did not form part of the printed specification, but was obtained in
electronic format directly from WIPO at
ftp.wipo.int/pub/published_pct_sequences

Sequence 1100 AA;

Query Match 15.9%; Score 541.5; DB 4; Length 1100;
Best Local Similarity 26.9%; Pred. No. 6.4e-43;
Matches 165; Conservative 90; Mismatches 223; Indels 135; Gaps 25;

QY 96 NNPDMILNKQVGVY---VDLF-----AGDLKGLKDKIPYQELGLTYLHMLPFX 145
 DB 11 NDPQWYKD-----ATYOVHVYSFYDANNDCIGDPAGLEKDYADAGVNTMLLPFY- 64
 QY 146 CPEGKSDGAYAVSYRDVNPALGTIGDLREVIALAHEAGISAIVDFIFNHTSNEHEW--- 202
 DB 65 -PSPRDGOYDIAQVRGHSVDGSLADARRFLAEAHRRGLRVITELVINHTSDQHPWFR 123
 QY 203 AQRCAAGDPLFDNFYIFPPDRRMPDQYRTLEIFPPDQHPGFSQLEDRRWYWT----- 256
 DB 124 ARHAKKGSRRAD--YYVMSD--SDEKYOGT-RIIFID-----TEGSNMTWMDPAQOY 170
 QY 257 ---TNSPQMDLNSNPNWPFVFRAMAGEMLFILANIGDILIRMDAVAFITWQMGSCENLPOA 313
 DB 171 YWHRFYSHOPDINFPNPOVLREVLDGVMRYWLDGVDGLRDLAIPLYLIEDGSSSENLPET 230
 QY 314 HALIFAFNAVMRIAPAVFEKSEALVHPDQVQYIG-----QDECOIGYNPLQWALLMN 367
 DB 221 HQVLKRIAEIDAHYPRDMLAEANQWPEDRPFYGGEDCGSGDCCHMAFHPPLMRMYM 290
 QY 368 TLATREVNILHQALTYRHNLPHTTAVVNYVRSNHDIGWTFADDEDAAYLGISGYDRQPLN 427
 DB 291 AIAQDRYPITDILKQTDIPANCMQALFLRNHDELTELEMTD-----DERDYI- 339
 QY 428 RFFVNRPFSGSFARVPFOVNPSTGDCVYS---GTAALVGLAQDDPHAVDRIKLILYSIAL 484
 DB 340 -----WNHYAADRRAKINIGIRRLAPLYERRR--RIELHSLILL 378
 QY 485 STGGLPLYLGDVEGT-----LINDD-----WSQDSN---KSDDSRWARRPYNEAL 528
 DB 379 SMPGPTLYYGGELIMGDMNITYGDRDGYRTPMQWGVDRNGGFSRADPALV-LPFLIDPL 437
 QY 529 Y-----AQRNDSTAGQTYOGLRHMIANVROSNPEDGRLTYFNTHNHIIGYI-- 578
 DB 438 YGYQTINVEAQARDHS---LIMMMRRLIAVRSOOKAFGRSLKMLAPSNRIILAYIRE 493
 QY 579 -----RNNALLAFNPESEYPOVTVAHTLQAMPEKAX-----DLGGKTV-----SLNDQL 623
 DB 494 YAEGRKDSITCVANILSRRAQNV-----ELDLASHAGKVPENMIGMSFPPIGELTYIL 547
 QY 624 TLQPYQVMMLEIA 636
 DB 548 TLPPYGFYWFYLA 560
 RESULT 14
 ABU38432
 ID ABU38432 standard; protein; 1100 AA.
 AC ABU38432;
 XX
 DT 19-JUN-2003 (first entry)
 XX
 DE Protein encoded by Prokaryotic essential gene #23959.
 XX
 KM Antisense; prokaryotic essential gene; cell proliferation; drug design.
 OS Pseudomonas aeruginosa.
 XX
 PN WO20027183-A2.
 XX
 PD 03-OCT-2002.
 XX
 PF 21-MAR-2002; 2002WO-US009107.
 XX
 PR 21-MAR-2001; 2001US-00815242.
 PR 06-SEP-2001; 2001US-00948993.
 PR 25-OCT-2001; 2001US-0342923B.
 PR 08-FEB-2002; 2002US-00072851.
 PR 06-MAR-2002; 2002US-0362699F.
 XX
 PA (ELIT-) ELITRA PHARM INC.
 XX

PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
 PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
 DR WPI; 2003-029926/02.
 DB N-PSDB; ACAA42302.
 PT New antisense nucleic acids, useful for identifying proteins or screening
 PT for homologous nucleic acids required for cellular proliferation to
 PT isolate candidate molecules for rational drug discovery programs.
 PS Claim 25; SEQ ID NO 66356; 1766bp; English.
 XX
 CC The invention relates to an isolated nucleic acid comprising any one of
 CC the 6213 antisense sequences given in the specification where expression
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:
 CC (1) a vector comprising a promoter operably linked to the nucleic acid
 CC encoding a polypeptide whose expression is inhibited by the antisense
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
 CC polypeptide or its fragment whose expression is inhibited by the
 CC antisense nucleic acid; (4) an antibody capable of specifically binding
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
 CC proliferation or the activity of a gene in an operon required for
 CC proliferation; (7) identifying a compound that influences the activity of
 CC the gene product or that has an activity against a biological pathway
 CC required for proliferation, or that inhibits cellular proliferation; (8)
 CC identifying a gene required for cellular proliferation or the biological
 CC pathway in which a proliferation-required gene or its gene product lies
 CC or a gene on which the test compound that inhibits proliferation of an
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
 CC compound's activity; (11) a culture comprising strains in which the gene
 CC product is overexpressed or underexpressed; (12) determining the extent
 CC to which each of the strains is present in a culture or collection of
 CC strings; or (13) identifying the target of a compound that inhibits the
 CC proliferation of an organism. The antisense nucleic acids are useful for
 CC identifying proteins or screening for homologous nucleic acids required
 CC for cellular proliferation to isolate candidate molecules for rational
 CC drug discovery programs, or for screening homologous nucleic acids
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
 CC the target prokaryotic essential genes. Note: The sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 1100 AA:
 Query Match 15.9%; Score 541.5; DB 6; Length 1100;
 Best Local Similarity 26.9%; Pred. No. 6; 4e-43;
 Matches 165; Conservative 90; Mismatches 223; Indels 135; Gaps 25;
 QY 96 NNPDMILNKQVGVY---VDLF-----AGDLKGLKDKIPYQELGLTYLHMLPFX 145
 DB 11 NDPQWYKD-----ATYOVHVYSFYDANNDCIGDPAGLEKDYADAGVNTMLLPFY- 64
 QY 146 CPEGKSDGAYAVSYRDVNPALGTIGDLREVIALAHEAGISAIVDFIFNHTSNEHEW--- 202
 DB 65 -PSPRDGOYDIAQVRGHSVDGSLADARRFLAEAHRRGLRVITELVINHTSDQHPWFR 123
 QY 203 AQRCAAGDPLFDNFYIFPPDRRMPDQYRTLEIFPPDQHPGFSQLEDRRWYWT----- 256
 DB 124 ARHAKKGSRRAD--YYVMSD--SDEKYOGT-RIIFID-----TEGSNMTWMDPAQOY 170
 QY 257 ---TNSPQMDLNSNPNWPFVFRAMAGEMLFILANIGDILIRMDAVAFITWQMGSCENLPOA 313
 DB 171 YWHRFYSHOPDINFPNPOVLREVLDGVMRYWLDGVDGLRDLAIPLYLIEDGSSSENLPET 230
 QY 314 HALIFAFNAVMRIAPAVFEKSEALVHPDQVQYIG-----QDECOIGYNPLQWALLMN 367
 DB 221 HQVLKRIAEIDAHYPRDMLAEANQWPEDRPFYGGEDCGSGDCCHMAFHPPLMRMYM 290
 QY 368 TLATREVNILHQALTYRHNLPHTTAVVNYVRSNHDIGWTFADDEDAAYLGISGYDRQPLN 427
 DB 291 AIAQDRYPITDILKQTDIPANCMQALFLRNHDELTELEMTD-----DERDYI- 339

```

QY 428 REFVNRFDGSFARGVFPQYNPSTGCRVS---GTAALVGLAODDGHANDRIKLYSTAL 484
DB 340 -----MNYAARDRANINIGIRRLAPLVERDR---RIELHSLDL 378
QY 485 STGGLPLIYLGDEVGT-----LNDLDD-----WSODSN---KSDSRMARPPRYNEAL 528
DB 379 SMRGPTLYYGDDELGMGNILYIGDDRDGVRTPMOWEVDNRGGSRADPAKLV-LPFLDLP 437
QY 529 Y-----AQRNDSTAAAGIYQGLRHMIAVRQSNPRPDGRLVTFTNNKHITGYI-- 578
DB 438 YGYQTINVEAQRADPHS-----LNNMRRLLAVRSQQKXFGSGSLMMLAPSNRIIAYIRE 493
QY 579 -----RNNALLAFGNPSEYPTVTAHTLOAMPFAH-----DLIGKTV-----SLNDL 623
DB 494 YABGRQDSILCVANLSRAQAV-----ELDLASHAGKPVEMTGMSPPIGELTYL 547
QY 624 TLQPYQVMLEIA 636
DB 548 TLPPYGFYWFYLA 560

RESULT 15
ABU21832
ID ABU21832 standard; protein; 1154 AA.
XX
XX ABU21832;
AC
XX 19-JUN-2003 (first entry)
DT
XX Protein encoded by Prokaryotic essential gene #7359.
DE
XX Antisense; prokaryotic essential gene; cell proliferation; drug design.
KM Burkholderia fungorum.
XX
XX WO200277183-A2.
XX
XX 03-OCT-2002.
PD
XX 21-MAR-2002; 2002WO-US009107.
XX
XX 21-MAR-2001; 2001US-00815242.
XX
XX 06-SEP-2001; 2001US-00948993.
XX
XX 25-OCT-2001; 2001US-0342923P.
XX
XX 08-FEB-2002; 2002US-00072851.
XX
XX 06-MAR-2002; 2002US-0362699P.
XX
XX (ELIT-) ELITRA PHARM INC.
PA
XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JM;
PI Wall D, Trawick JD, Cair GJ, Yamamoto R, Forsyth RA, Xu HH;
PI WPI; 2003-029926/02.
XX
XX N-PSDB; ACA25702.
XX
XX New antisense nucleic acids, useful for identifying proteins or screening
PT for homologous nucleic acids required for cellular proliferation to
PT isolate candidate molecules for rational drug discovery programs.
XX
XX Claim 25; SEQ ID NO 49756; 1766BP; English.
XX
XX The invention relates to an isolated nucleic acid comprising any one of
XX the 623 antisense sequences given in the specification where expression
XX of the nucleic acid inhibits proliferation of a cell. Also included are:
XX (1) a vector comprising a promoter operably linked to the nucleic acid
XX encoding a polypeptide whose expression is inhibited by the antisense
XX nucleic acid; (2) a host cell containing the vector; (3) an isolated
XX polypeptide or its fragment whose expression is inhibited by the
XX antisense nucleic acid; (4) an antibody capable of specifically binding
XX the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
XX proliferation or the activity of a gene in an operon required for
XX proliferation; (7) identifying a compound that influences the activity of

```

```

CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than S. aureus, S. typhimurium,
CC K. pneumoniae or P. aeruginosa. The present sequence is encoded by one of
CC the target prokaryotic essential genes. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 1154 AA;
SQ

```

```

Query Match 15.9%; Score 541; DB 6; Length 1154;
Best Local Similarity 27.6%; Pred. No. 7.8e-43;
Matches 155; Conservative 86; Mismatches 224; Indels 96; Gaps 18;

```

```

QY 118 GDLKGLKDKIPYQDELGLTYLHMLPFCPEKSGDGYAVSSYRDVNPALGTIGLREVI 177
DB 65 GDFPLGLAKLDYIAELGVNAIWLIPFY--PSRRDDGDYIDYRVVHDYGLSDVKKRFI 122
QY 178 AALHAGISAVYDFIFNTSNEHEWACCAAGDPLFDN-FYIIFPRRMPDQYDTLREI 236
DB 123 QEAHARGIRVITTELVINTSDHPFQARAKPESNERNYVWSN---TDQKYLETRI 179
QY 237 FPDGHPGGSQLED---GRWWTTFNSFQMDLNTSNPWPFPAMAGEMFLANLGYDIIRM 293
DB 180 FIDSEPSWT--HDPVAGAYYWHFYSHQPLNPNPVALAEVLQVWFMDMGIDGRL 237
QY 294 DAVAFIWKQMGTSCEINLPQAHALIRAFNAVRILAAVAFYKSEALIVPDQVQYIG-ODE 352
DB 238 DAVPYLVEREGTNNENLPETHAVLKIATIDAEYFNMMLAEANQWEDVEYEGDDE 297
QY 353 QGIGNPLQMLLNNMTLATREVNLLHQALTYRHNPEHTAVWNVYRSHDDIGWTF-ADED 411
DB 298 CHMAHFPLMPRIYMSIASEDRFPITDTRQTPDLAETNQAIFLRNDELILEWVTJSE 357
QY 412 AAYL-GISGYDHRQFINRFVNRFDGSFARGVFPQYNPSTGDCRVSGTAALVGLAODD 470
DB 358 RDYLMNTYASDRARLN-----LGRRLAPLMEBRDR 389
QY 471 HAVDRIKLYSLAISTGGLPLIYLGDEVGT-----LNDLDD-----WSODSN---KSD 515
DB 390 R---RIELINSLILSMPTPIVYIGDEIGMDNIIHIGRDRVCRPMQSSSRNGFSRAD 446
QY 516 SRMARPPRYNEALY-----AQRNDSTAAAGIYQGLRHMIAVRQSNPRPDGRLVTFT 567
DB 447 PEGLVLPFVMSIVGFDAVNVFAQSRDPS-----LNNMRRLLAVRSQQKXFGSGSLM 502
QY 568 NTNNKHITGYIYR---NNALLAFGNFSEYPTVTAHTLOAMPFAHDLIGKTVSLND- 622
DB 503 KPNRKILAVIREMGEPPILCVANLSRAP-----QAVELDSEPDGVSPIEMTADS 554
QY 623 -----TLQPYQVM 632
DB 555 VPPPIGLTYLTLPPYGFYLA 575

```

Search completed: May 25, 2004, 19:30:26
Job time : 66 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using SW model

Run on: May 25, 2004, 19:22:06 / Search time 45 Seconds
(without alignments)
4459.324 Million cell updates/sec

Title: US-09-807-146-1

Perfect score: 3408
Sequence: 1 MLTPQGVLLIQYLKTRL.....VSLNDLTPQYVMMLEIA 636

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: SPTREMBL_25:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phage:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp Vertebrate:*
- 14: sp Unclassified:*
- 15: sp_virus:*
- 16: sp_bacteriaph:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3408	100.0	636	2 Q9ZEU2	Q9ZEU2 neisseria p
2	3338	97.9	636	2 Q84HD6	Q84HD6 neisseria m
3	3333	97.8	636	2 Q84HD5	Q84HD5 neisseria p
4	1646	48.3	701	16 Q7UG17	Q7UG17 rhodospirillum rubrum
5	1319.5	38.7	644	16 Q9RVT9	Q9RVT9 deinococcus
6	1126.5	33.1	644	16 Q8RGX2	Q8RGX2 xanthomonas
7	1116.5	32.8	637	16 Q8PS12	Q8PS12 xanthomonas
8	919.5	27.0	584	16 Q9A959	Q9A959 caulobacter
9	550	16.1	1106	16 Q88FNO	Q88FNO pseudomonas
10	548	16.1	1160	16 Q88FNO	Q88FNO pseudomonas
11	541.5	15.9	1100	16 Q911W3	Q911W3 pseudomonas
12	541.5	15.9	1113	16 Q7W1R8	Q7W1R8 bordetella
13	538.5	15.8	1113	16 Q7W1R8	Q7W1R8 bordetella
14	529.5	15.5	572	16 Q54207	Q54207 streptomyces
15	529.5	15.5	572	16 Q82JF2	Q82JF2 streptomyces
16	529.5	15.5	598	16 Q8NNM1	Q8NNM1 corynebacter

17	529	15.5	1116	16 Q8PE49	Q8PE49 xanthomonas
18	525.5	15.4	1108	16 Q8KAR7	Q8KAR7 chlorobium
19	523.5	15.4	612	16 Q8FND9	Q8FND9 corynebacter
20	522	15.3	1098	16 Q89FD4	Q89FD4 bradyrhizob
21	522	15.3	1116	16 Q8PR14	Q8PR14 xanthomonas
22	521.5	15.3	1108	16 Q881X1	Q881X1 pseudomonas
23	520.5	15.3	552	16 Q9RST7	Q9RST7 deinococcus
24	520	15.3	965	2 Q7WU15	Q7WU15 thermus the
25	517.5	15.2	668	17 Q8TOA8	Q8TOA8 methanococcus
26	516	15.1	566	2 Q54203	Q54203 streptomyces
27	515.5	15.1	566	16 Q9L1K3	Q9L1K3 streptomyces
28	515	15.1	965	2 Q9RA59	Q9RA59 thermus cal
29	511.5	15.0	583	16 Q82SR1	Q82SR1 streptomyces
30	509.5	15.0	601	16 Q07176	Q07176 mycobacteri
31	509.5	15.0	601	16 Q7U2S8	Q7U2S8 mycobacteri
32	484	14.2	585	2 Q7WU14	Q7WU14 actinoplanes
33	467	13.7	553	16 Q7U4V8	Q7U4V8 synecchococcus
34	452.5	13.3	568	16 Q8FHS2	Q8FHS2 escherichia
35	450.5	13.2	571	16 Q7U1S9	Q7U1S9 rhodospirillum
36	445	13.1	544	16 Q92X63	Q92X63 rhizobium m
37	404.5	11.9	585	16 Q7V1B8	Q7V1B8 proclostridium
38	404	11.9	565	16 Q89VZ1	Q89VZ1 bradyrhizob
39	393	11.5	566	16 Q926R8	Q926R8 listeria in
40	386.5	11.3	434	16 Q8X816	Q8X816 escherichia
41	373	10.9	560	16 Q8Y3U6	Q8Y3U6 listeria mo
42	372	10.9	547	16 Q88ZP5	Q88ZP5 lactobacilli
43	349	10.2	575	16 Q88ZX0	Q88ZX0 lactobacilli
44	345.5	10.1	575	16 Q7U3J7	Q7U3J7 synecchococcus
45	342	10.0	537	16 Q835M8	Q835M8 enterococcus

ALIGNMENTS

RESULT 1

Q9ZEU2 PRELIMINARY; PRT; 636 AA.

AC Q9ZEU2; 01-MAY-1999 (TREMBLrel. 10, Created)

DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)

DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)

DE Amylosucrase (EC 2.4.1.4) (Fragment).

GN AMYLOSUCRASE.

OS Neisseria polysaccharaea.

OC Bacteria: Proteobacteria: Betaproteobacteria: Neisseriales;

OC Neisseriaceae; Neisseria.

OX NCBI_TaxID=489;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=ATCC 43768;

RX MEDLINE=99102197; PubMed=9882648;

RA Potocki de Montalk G., Remaud-simeon M., Willemot R.M., Plancho V.,

RA Monnan P.;

RT "Sequence analysis of the gene encoding amylosucrase from Neisseria

RT polysaccharaea and characterization of the recombinant enzyme.";

RL J. Bacteriol. 181:375-381(1999).

DR EMBL: AJ011781; CAA09772.1; --

DR PDB: 1G5A; 07-NOV-01.

DR PDB: 1UG9; 31-OCT-01.

DR PDB: 1UG1; 31-OCT-01.

DR PDB: 1MY7; 18-DEC-02.

DR PDB: 1MW0; 18-DEC-02.

DR PDB: 1MW1; 18-DEC-02.

DR PDB: 1MW2; 18-DEC-02.

DR PDB: 1MW3; 18-DEC-02.

DR GO: GO:0004556; F:alpha-amylase activity; IEA.

DR GO: GO:0016757; F:transferrase activity; transferring glycosyl. . .; IEA.

DR GO: GO:0005975; P:carbohydrate metabolism; IEA.

DR InterPro: IPR006047; Alpha-amyl cat.

DR Pfam: PF00128; alpha-amylase; 1.

DR GlycoSyl:Transferrase; Transferrase.

FT NON_TER 1

SEQUENCE 636 AA; 72343 MW; B7656C19BF1A2065 CRC64;

RX PubMed=12517860;
 RA Zhu P., Tsang R.S.W., Tsai C.M.;
 RT "Nonencapsulated *Neisseria meningitidis* Strain Produces Amylopectin
 from Sucrose: Altering the Concept for Differentiation between *N.*
 RT meningitidis and *N. polysacchara*."
 RL J. Clin. Microbiol. 41:273-278(2003).
 DR EMBL; AY099335; AAMS1153.1; -
 DR GO; GO:0004556; F:alpha-amylase activity; IEA.
 DR InterPro; IPR006047; Alpha_amyl_cat.
 DR InterPro; IPR006589; Alp_amyl_cat_sub.
 DR Pfam; PF00128; alpha-amylase; 1.
 DR SMART; SM00642; Amy; 1.
 SQ SEQUENCE 636 AA; 72271 MW; 98F81B0BFAC39A0C CRC64;

Query Match 97.8%; Score 3333; DB 2; Length 636;
 Best Local Similarity 97.5%; Pred. No. 1.9e-259;
 Matches 620; Conservative 9; Mismatches 7; Indels 0; Gaps 0;

1 MTTPTQVGLITIOYKTRITIDITPBRAGIEKSEEDRFRSRMDTHFPKLNELDSVYG 60
 1 MTPPTQVGLITIOYKTRITIDITPBRAGIEKSEEDRFRSRMDTHFPKLNELDSVYG 60
 61 NNEALLPMLLEMLLAQAMQSYQSNSSLKDIARENNPMLSNKQVGVGYVDLFAGDL 120
 61 NNEALLPMLLEMLLAQAMQSYQSNSSLKDIARENNPMLSNKQVGVGYVDLFAGDL 120
 121 KKLKDKIPYFOELGLTYLHMLPKFCPEKSGDGVAVSSYRDVNPALGTTGDLREYIAL 180
 121 KKLKDKIPYFOELGLTYLHMLPKFCPEKSGDGVAVSSYRDVNPALGTTGDLREYIAL 180
 181 HBAAGSAVVDPIFNHTSNEHEWAQRCAGDPLFDNYYTIPDRPMDOYDRLRETFPQ 240
 181 HBAAGSAVVDPIFNHTSNEHEWAQRCAGDPLFDNYYTIPDRPMDOYDRLRETFPQ 240
 241 HEGGSQLEDGKRWVWTFNSFQMDLNSNPWFRAAGEMFLANLGVUILMDAVAFIM 300
 241 HEGGSQLEDGKRWVWTFNSFQMDLNSNPWFRAAGEMFLANLGVUILMDAVAFIM 300
 301 KOMGTSCEMLPOAHALIRAFNAVMRIAADAVFEKSAIVHPQVVOYIGODECOIGYNPL 360
 301 KOMGTSCEMLPOAHALIRAFNAVMRIAADAVFEKSAIVHPQVVOYIGODECOIGYNPL 360
 361 QMALLMNTLATREVNILHQAITYRHNLPHBTAMVNVYRSHDIMGTFADDAAYIGISGY 420
 361 QMALLMNTLATREVNILHQAITYRHNLPHBTAMVNVYRSHDIMGTFADDAAYIGISGY 420
 421 DHRQFLNRFVVRPDSFARGVPFOYNPTGDCRVSGTAAALVGLAQDDPHAVRIKILY 480
 421 DHRQFLNRFVVRPDSFARGVPFOYNPTGDCRVSGTAAALVGLAQDDPHAVRIKILY 480
 481 STALSTGSLPLIYLGDVGTLDNDWSQDSNKSDDSRRAHRRPRYNEALYAQRNDPSTAG 540
 481 STALSTGSLPLIYLGDVGTLDNDWSQDSNKSDDSRRAHRRPRYNEALYAQRNDPSTAG 540
 541 QTYOGRHMIAROSRPERDGGRLVFNNTNKHIIYIRNNALLAFGNFSEYPTQTAHT 600
 541 QTYOGRHMIAROSRPERDGGRLVFNNTNKHIIYIRNNALLAFGNFSEYPTQTAHT 600
 601 LQAMPKADHLIGKTVSLNODLTIQPYVMWLEIA 636
 601 LQAMPKADHLIGKTVSLNODLTIQPYVMWLEIA 636

RESULT 4
 ID 07UG17 PRELIMINARY; PRT; 701 AA.
 AC 07UG17;
 DT 01-OCT-2003 (TEMBLrel. 25, Created)
 DT 01-OCT-2003 (TEMBLrel. 25, Last sequence update)
 DT 01-OCT-2003 (TEMBLrel. 25, Last annotation update)
 DE Alpha-amylase, amylase (EC 2.4.1.4).
 GN RB5196.

OS Rhodopirellula baltica.
 OC Bacteria; Planctomycetes; Planctomycetacia; Planctomycetales;
 OC Planctomycetaceae; Pirellula.
 OK NCBI_Taxid=117;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=1;
 RX MEDLINE=22735913; PubMed=12835416;
 RA Gloeckner F.O., Kube M., Bauer M., Teeling H., Lombardot T.,
 RA Ludwig W., Gade D., Beck A., Bozzym K., Heilmann K., Rabus R.,
 RA Schlesner H., Aumann R., Reinhardt R.;
 RT "Complete genome sequence of the marine planctomycete Pirellula sp.
 RT strain 1."
 RL Proc. Natl. Acad. Sci. U.S.A. 100:8298-8303(2003).
 DR EMBL; BX294141; CAD78342.1; -
 DR Glycoyltransferase; Transferase; Complete proteome.
 SQ SEQUENCE 701 AA; 81070 MW; 29FF2F85C45F454 CRC64;

Query Match 48.3%; Score 1646; DB 16; Length 701;
 Best Local Similarity 52.1%; Pred. No. 1.5e-123;
 Matches 316; Conservative 94; Mismatches 180; Indels 16; Gaps 6;

43 RMDTHFPKLNELDSVYGNNALLPMLLEMLLAQAMQSYQSNSSLKDIARENNPMDL 102
 94 RIVGMDLFGYLLHLYGQWDFYHLBQILITIRGMRBEBQLKODEHRINNEBYQ 153
 103 SNKQGVGYVDLFGAGDLKGLDKIPYFOELGLTYLHMLPKFCPEKSGDGVAVSSYRD 162
 154 SEKLVGALYVLFSENLGELRKQIPYFODLGLSYHMLPLAVRPGNDGVALSNYS 213
 163 VNPALGTTGDLREYIALHBAAGSAVVDPIFNHTSNEHEWAQRCAGDPLFDNYYTIPD 222
 214 VDPRTGTTDLKLLADDLREAGILVDFVFHTADHDHMAQAGSNEGYKYTFPD 273
 223 RMPDQYRTLEIFPDQHPGFSQLEDG-RWWTTFNSFQMDLNSNPWFRAAGEMFLANL 280
 274 REVPOYERTLEIFPTVARGNFT-WHDQMCMVWTFNSFQMDLNSNPWFRAAGEMFLANL 332
 281 LFLANLGVUILMDAVAFIMKOMGTSCEMLPOAHALIRAFNAVMRIAADAVFEKSAIVH 340
 333 LFLANLGVUILMDAVAFIMKOMGTSCEMLPOAHALIRAFNAVMRIAADAVFEKSAIVH 392
 341 PPOVVOYIGODECOIGYNPLQMALLMNTLATREVNILHQAITYRHNLPHBTAMVNVYRSH 400
 393 PPOVVOYIGODECOIGYNPLQMALLMNTLATREVNILHQAITYRHNLPHBTAMVNVYRSH 452
 401 DDIGTFADDAAYIGISGYDHRQFLNRFVVRPDSFARGVPFOYNPTGDCRVSGTAA 460
 453 DDIGTFADDAAYIGISGYDHRQFLNRFVVRPDSFARGVPFOYNPTGDCRVSGTAA 512
 461 ALVGLAQ-----D-PAVVRILKLYSLSTGSLPLIYLGDVGTLDNDWSQDSNK 512
 513 SLAAGSQALESRDWKELAIRILNMMMLSVGGIPLYTIGEEGALANDVDVYDPAK 572
 513 SDDSRRAHRRPRYNEALYAQRNDPSTAGQIYOGRLMIAROSRPERDGGRLVFNNTNKH 572
 573 AGDTMVRHRPRKQWKFLELDGDSIRGRITYSIKRLIARMTTAFGLDMLLLEBP 632
 573 HIIGITR--NNALLAFGNFSEYPTQTAHTIQ-AMPKADHLIGKTVSLNODLTIQ 627
 633 HIIGITR--NNALLAFGNFSEYPTQTAHTIQ-AMPKADHLIGKTVSLNODLTIQ 692
 628 YQVWML 633
 693 YEFMWL 698

RESULT 5
 ID 09RV79 PRELIMINARY; PRT; 644 AA.
 AC 09RV79;
 DT 01-MAY-2000 (TEMBLrel. 13, Created)
 DT 01-MAY-2000 (TEMBLrel. 13, Last sequence update)
 GN RB5196.

DT 01-JUN-2003 (TREMBLREL. 24, last annotation update)
 GN Alpha-AMYLASE.
 OS Deinococcus radiodurans.
 OC Bacteria; Deinococcus-Thermus; Deinococci; Deinococcales;
 OC Deinococcaceae; Deinococcus.
 NCBI_Taxid=1259;
 RN [1]
 RC SEQUENCE FROM N.A.
 RP STRAIN=R1 / ATCC 13939 / DSM 20539 / NCIB 9279;
 RX MEDLINE=20036896; PubMed=10567266;
 RA White O., Eisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,
 RA Moffat K.S., Qin H., Jiang L., Pamphile W., Crosby W., Shen M.,
 RA Vamathevan J.J., Lam P., McDonald L., Utechtack T., Zaleski C.,
 RA Makarova K.S., Aravind L., Daly M.J., Minton K.W., Fleischmann R.D.,
 RA Ketchum K.A., Nelson K.E., Salzberg S., Smith H.O., Venter J.C.,
 RA Fraser C.M.;
 RT "Genome sequence of the radioresistant bacterium Deinococcus
 radiodurans R1.";
 RL Science 286:1571-1577(1999).
 DR EMBL; AE001946; AAF10510.1; -.
 DR PIR; C75457; C75457.
 DR TIGR; DR0933; -.
 DR GO; GO:0004556; F:alpha-amylase activity; IEA.
 DR GO; GO:0016829; F:lyase activity; IEA.
 DR GO; GO:0005975; P:carbohydrate metabolism; IEA.
 DR InterPro; IPR006047; Alpha_amyl_cat.
 DR Pfam; PF00128; alpha-amylase; 1.
 DR Lyase; Complete proteome.
 SQ SEQUENCE 644 AA; 71668 MW; 610077561181978 CRC64;

Query Match 38.7%; Score 1319.5; DB 16; Length 644;
 Best Local Similarity 42.1%; Pred. No. 2,6e-97;
 Matches 281; Conservative 110; Mismatches 215; Indels 61; Gaps 14;

1 MLTPPOQGLIIOYKTRIDITYTBOAGIEKSDMPOFSRMDTHEPKMNEIDSYG 60
 1 MLTPD-----LAARL-----RAFDDBDAETFRILRYGELADNIRAYG 43
 61 NN-EALL-PMLEMLLAQAWOSYQORNSIKIDIRENNPDILSNKQGVGYVDLPAG 118
 44 NHADALIGLLEVM-----HAYHARPADIKRIDERLRPDLQGPBWGYAYVDFRAG 99
 119 DIKGLKDKIPYFOELGLTYLHMLPKCEBGSQGVAVSYRDYNPALGTTGIDREYIA 178
 100 TRGVGERLEYLEGLGVTYLHLLPLRPDGDNDGAYVQDYRSVRPDIGTIDDLASLAR 159
 179 ALHEGISAIVDPFINHTSNEHEMQRCAQDPLDNFYIFPPDRMPQVORTLREIF? 238
 160 ELRGGISLVDLVNHNVAEHEHMAVRATGAARVDYFHPIDPTQPDVARTLPEIF? 219
 239 DQHPGFS-QLEDGHWVWTTFNSFQMDLNSNPWFRRMAGHMLFLANIGVDILMDVA 297
 220 DPAENFTWNGEAGGWMVTTFRSQMDVWNGNPAIFREYRLDILTLARGVEVFFLDIA 279
 298 FTKWQMGTSCEMLPQAHALIRAFNAVRIAPAVPFKSAIYHPQOVVOYIGQD----- 352
 280 FLMKRLGTCQNOPEVHRLTHALRAATRIAPAVAFKAAIYAPGGLIHYLGSRHHGV 339
 353 CQIGNEPLQMLMNTLAREVNLHQALTYRHNIPETHAMVNYRSHDQWTFADEDA 412
 340 SPMAHNSIMVQIMSSIASRDTLLAALAAPPKPKTNTTWGVYRCHDIDMALADEDA 399
 413 AVTIGSGYHROFLNRFNVRFDGAFARVPQYNPSTGDCRVSGTAAVLGL----- 465
 400 AAVGSLGPAHRLFLSFYSGEPFGSFAKGLVQHPHQTDRILSTGASLAGLDLALENG 459
 466 -AQDDPHAVDRIKLYSIALSTGGLPLTYLGDVGTLLDDWSQDSNKSDDSRMAHRRPY 524
 460 DAERVNDALARLLLAHVAVMGFGVPLTYMGDELALINDTDFAAVPAHAADNRWHRPQY 519
 525 NNAIVAQ-----RNDPSFAGQIYQGLHMTIAVQNSNRPFGGRVLTENTNKKH----- 574

DB 520 DWELVASHQADATGQPTPAGRFAGRLHRLARRTPHLLA-----STESPILSPD 573
 QY 575 --IGYIR-----NNALLAFNPFSEYPOVTATLQ-AMFKRADLIGKTVSL-NDQTLQ 626
 DB 574 PCVLLIREHRTGVLQVYNSEHHITFPTRPDEQGLAVAHDLSSQPHLGDDLAL 633
 QY 627 PYQVWML 633
 DB 634 FYRALML 640

RESULT 6
 Q8PGX2 PRELIMINARY; PRT; 644 AA.
 AC Q8PGX2;
 DT 01-OCT-2002 (TREMBLREL. 22, Created)
 DT 01-OCT-2002 (TREMBLREL. 22, Last sequence update)
 DT 01-JUN-2003 (TREMBLREL. 24, Last annotation update)
 DS Amylase or alpha amylase.
 GN XAC3490.
 OS Xanthomonas axonopodis (pv. citri).
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
 OC Xanthomonadaceae; Xanthomonas.
 NCBI_Taxid=92829;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=306 / ATCC 13902 / XV 101;
 RX MEDLINE=22022145; PubMed=12024217;
 RA da Silva A.C.R., Ferro J.A., Reinach F.C., Farah C.S., Furlan L.R.,
 RA Queglio R.B., Monteiro-Vitorino C.B., Van Sluys M.A., Almeida N.F.,
 RA Alves L.M.C., do Amaral A.M., Bertolini M.C., Camargo L.E.A.,
 RA Camarotte G., Camarvan F., Cardoso J., Chabergo F., Ciapina L.P.,
 RA Claretelli R.M.B., Coutinho L.V., Cursino-Santos J.R., El-Dotry H.,
 RA Faria J.B., Ferreira A.J.S., Ferreira R.C.C., Ferro M.I.T.,
 RA Formiglieri E.F., Franco M.C., Greggio C.C., Gruber A.,
 RA Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,
 RA Locelli E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.,
 RA Martins E.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,
 RA Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,
 RA Pereira H.A., Rossi A., Sena J.A.D., Silva C., de Souza R.F.,
 RA Spindola L.A.F., Takita M.A., Tamura R.E., Teixeira E.C., Tezra R.I.D.,
 RA Trindade dos Santos M., Truffi D., Teai S.M., White P.F.,
 RA Serubai J.C., Kitajima J.P.;
 RT "Comparison of the genomes of two Xanthomonas pathogens with differing
 host specificities.";
 RL Nature 417:459-463(2002).
 DR EMBL; AE011997; AAM38333.1; -.
 DR GO; GO:0004556; F:alpha-amylase activity; IEA.
 DR GO; GO:0005975; P:carbohydrate metabolism; IEA.
 DR InterPro; IPR006047; Alpha_amyl_cat.
 DR Pfam; PF00128; alpha-amylase; 1.
 DR Complete proteome.
 SQ SEQUENCE 644 AA; 70470 MW; 041ED026FE935CCD CRC64;

Query Match 33.1%; Score 1126.5; DB 16; Length 644;
 Best Local Similarity 39.0%; Pred. No. 9e-82; Indels 33; Gaps 10;
 Matches 240; Conservative 107; Mismatches 236;

43 EMDTFPKLMEELSVYGNNEALLPMLLEMLLAQAWOSYQORNSIKIDIRENNPDWL 102
 30 RYDQASRLDLALHLYQGRADYASWLQWGEVGDIAQRQALQTDSTR--HAGWFG 87
 103 SNKQGVGYVDLFAFGDLKGLDKTIPYFOELGLTYLHMLPKCEBGSQGVAVSYRD 162
 88 QQHMIGYAVYDREFAGTQGVARVPYFOELGVRLHLLPLRARAGNDGFAVSDGQ 147
 163 VNPALGTIGDLREVIALAHEGISAIVDPFINHTSNEHEMQRCAQDPLFNFYIIPD 222
 148 VEPISGSDNDVALTSRLREKGISLCADFVNLHNTDDMAQAAAGARVLDYHHFAD 207
 223 RMPDQVORTLREIFPDHPGFSQLED-GRVWTTFNSFQMDLNSNPWFRRMAGHML 281

```

Db 208 RTLPDRYEATLGQVFPHTAPGNFTWVDITAQMMWTFYPCQDLNWSNPAVFGDMALMTL 267
Qy 282 FLAANGVDILRMDAFAITKMGKSCENLPQAHALIRAFNAMYRLAAPVFKSEATVAP 341
Db 268 RLANIGVEFRILDSATYIMKRGITGCMQPEAHTLLVALRAVDIVAPVVKAAEATVPM 327
Qy 342 DVGVGYIGQ-----DECOIGYNPLQMALMTLATREVVLLHQALTYRNLPEHTAMVY 396
Db 328 TOLPPYFGSGVDEGHECHLAHYSTLMAAGMSALAQROGDIILANVLAHSPPLRHCAMLSY 387
Qy 397 VRSDDIGMTFADDEDAAYIGSGYHROFLNRFVNRREDGSPARGVPFQYNSTGD--CR 454
Db 388 VARCHDIDGNVVLQHEACGNAAQPPSLRVDARFANAVGSYARGESFG---SSGDGVHG 444
Qy 455 VSGTAAALVGL--AODDPH-----AVDRIKLLYSIALSTGGPLIYLGDEVGTLDDDM 506
Db 445 TNGMAAALAGIQAAQEGADAAALAAVAVDLVLLYALMALMPGVPLIYMGDELAMVNDPXY 504
Qy 507 SODSNKSDSRWAHPRVNEALYAQRNDPSTAAGQIYOGIRHMIIVROSNPRFDGR--LV 565
Db 505 RHDPHROHEGRWHLRPPAMDWOLAAQRHDANSISGTYRRLKGLIKORPALGALADQALA 564
Qy 566 TETNNKHLIIGYIRNNALIAFGNFSRYPQTV-----TAHTLOAMPFYAHDLIGKTVS 618
Db 565 STALNDPRYFVLTGRGSPFLAHNFSDQPLDVELAALIGDGTLLAI----DDALGGAAAR 620
Qy 619 INODITLOPYQVMLE 634
Db 621 GDGSIPLPYGVVRLQ 636

RESULT 7
QBP512 PRELIMINARY; PRT; 637 AA.
AC QBP512;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, last annotation update)
DE 01-JUN-2003 (TrEMBLrel. 24, last annotation update)
DE Amylase or alpha amylase.
GN XCC3359.
OS Xanthomonas campestris (pv. campestris).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC Xanthomonadaceae; Xanthomonas.
OX NCBI_TaxID=340;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 33913 / NCPPB 528;
RX MEDLINE=2022145; PubMed=12024217;
RA da Silva A.C.R., Ferro J.A., Reinach F.C., Farah C.S., Furlan L.R.,
RA Quaggio R.B., Monteiro-Vitorello C.B., Van Sluys M.A., Almeida N.F.,
RA Alves L.M.C., do Amaral A.M., Bertolini M.C., Camargo L.B.A.,
RA Camarotte G., Camavan F., Cardozo J., Chambergio F., Chapina L.P.,
RA Ciccarelli R.M.B., Coutinho L.L., Cristiano-Santos J.R., El-Dorri H.,
RA Formighieri E.F., Franco M.C., Greggio C.C., Pedro M.I.T.,
RA Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,
RA Locali E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rosel N.M.,
RA Marcini E.C., Melanlis U., Menck C.F.M., Miyaki C.Y., Moon D.H.,
RA Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,
RA Pereira H.A., Rosel A., Sena J.A.D., Silva C., de Souza R.F.,
RA Spindola L.A.F., Takita M.A., Tamura R.E., Teixeira E.C., Tezza R.I.D.,
RA Trindade dos Santos M., Truffi D., Tsai S.M., White F.F.,
RA Setubal J.C., Kitajima J.P.;
RT "Comparison of the genomes of two Xanthomonas pathogens with differing
RT host specificities."
RT Nature 417:459-463 (2002).
DR EMBL; AB014454; AAM42629.1; -.
DR GO; GO:0004556; P:alpha-amylase activity; IEA.
DR GO; GO:0005975; P:carbohydrate metabolism; IEA.
DR InterPro; IPR006047; Alpha_amyl_cat.
DR Pfam; PF00128; alpha-amylase; 1.
KW Complete proteome.
SQ SEQUENCE 637 AA; 69295 MW; B42C5BF7B94ADFAC CRC64;

```

```

Query Match 32.8%; Score 116.5; DB 16; Length 637;
Best Local Similarity 39.3%; Pred. No. 5.6e-81;
Matches 241; Conservative 106; Mismatches 236; Indels 31; Gaps 10;

Qy 43 RMDTHFPLKMLNDSVYGNNEALLPMLMLLAQAMQSYORSNSLKIDIDIAENNPWDL 102
Db 31 RPDTHAPRLDLASTLYGDHADYVTWLPQWISTALGVVAQAPPAALRHLDSSRA--PGWFG 88
Qy 103 SNKQYGVGYVDLPAGDKLGLKDKTIFYOEGLTYTHMLPFLKCEGKSGKSGYAVSSYND 162
Db 89 QODMLGYAAYDRFAGTLRGVAERVPYLOELGVRYLHLPLFRARAGDDGPAVSDYQ 148
Qy 163 VNPALGTIGDLREVAALHEAGISAVVDYFIENHTSNEHEMAORCAAGDLPFNFYIYFD 222
Db 149 VEPALGSDDLVALTARKLRANISLCADPYLNTHTDDHAKMAQAAAGDRYLDYHHFND 208
Qy 223 RMPDQYRTLEIRFPDHPGGFQLEDDR--VWVTFNSFQMDLVNSPFWPRAAGEM 281
Db 209 RNAPDQYDTLVQVFPQAPGNFTWDETRQMMWTFYPCQDLNWSNPAVFGEMALMTL 268
Qy 282 FLAANGVDILRMDAFAITKMGKSCENLPQAHALIRAFNAMYRLAAPVFKSEATVAP 341
Db 269 RLANIGVEFRILDSATYIMKRGITGCMQPEAHTLLVALRAVDIVAPVVKAAEATVPM 328
Qy 342 DVGVGYIGQ-----DECOIGYNPLQMALMTLATREVVLLHQALTYRNLPEHTAMVY 396
Db 328 TOLPPYFGSGVDEGHECHLAHYSTLMAAGMSALAQROGDIILANVLAHSPPLRHCAMLSY 387
Qy 397 VRSDDIGMTFADDEDAAYIGSGYHROFLNRFVNRREDGSPARGVPFQYNSTGD--CR 454
Db 388 VARCHDIDGNVVLQHEAAGTAQPPSLRVDARFANAVGSYARGESFG---SSGDGVHG 445
Qy 455 VSGTAAALVGLAODDPH-----AVDRIKLLYSIALSTGGPLIYLGDEVGTLDDDM 506
Db 445 TNGMAAALAGVQAAHEHADAAAARVAVDLVLLYAVSLMPGVPLIYMGDELAPMDTX 505
Qy 507 SODSNKSDSRWAHPRVNEALYAQRNDPSTAAGQIYOGIRHMIIVROSNPRFDGR--LV 565
Db 506 LDDAQRHGRWHLRPPAMWELAAQRHDASTLAGVYTRLRALIRKAGLPALAAQTSGS 565
Qy 566 TETNNKHLIIGYIRNNALIAFGNFSRYP---QYTAHTLOAMPFYAHDLIGKTVS 620
Db 566 SVALGDARLFLTRGSPFLAHNFSDVPLVDTQGHALMV---LDTGTGAPBEH 621
Qy 621 QDLTLOPYQVMLE 634
Db 622 TELLPAYGVVRLQ 635

RESULT 8
Q9A959 PRELIMINARY; PRT; 584 AA.
AC Q9A959;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, last annotation update)
DE 01-JUN-2003 (TrEMBLrel. 24, last annotation update)
DE Amylase.
GN CCL135.
OS Caulobacter crescentus.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Caulobacteriales;
OC Caulobacteraceae; Caulobacter.
OX NCBI_TaxID=155892;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 19089 / CB15.
RX MEDLINE=21173698; PubMed=11259647;
RA Nierman W.C., Feldblyum T.V., Laub M.T., Paulsen I.T., Nelson K.E.,
RA Eisen J., Heidelberg J.F., Alley M.R.K., Ohta N., Maddock J.R.,
RA Potocka I., Nelson W.C., Newton A., Stephens C., Phadke N.D., Ely B.,
RA DeBoy R.T., Dodson R.J., Durkin A.S., Gwinn M.L., Haft D.H.,
RA Kolonay J.F., Smit J., Craven M.B., Khouli H., Shetty J., Berry K.,
RA Ueberlack T., Tran K., Wolf A., Vamathevan J., Ermolaeva M., White O.,

```

RA Salberg S.L., Venter J.C., Shapiro L., Fraser C.M.;
 RT "Complete genome sequence of *Caulobacter crescentus*,"
 RL Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).
 DR EMBL: AB005791; AK23119.1; --
 DR PIR: C67390; C67390.
 DR TIGR: CC1135; --
 DR GO: GO:0004556; P:alpha-amylase activity; IEA.
 DR GO: GO:0005975; P:carbohydrate metabolism; IEA.
 DR InterPro: IPR006047; Alpha.amyl.cat.
 DR Pfam: PF00128; alpha-amylase; 1.
 DR Complete proteome.
 KW SEQUENCE 584 AA; 63232 MW; D2788FE75ACD624 CRC64;

Query Match 27.0%; Score 919.5; DB 16; Length 584;
 Best Local Similarity 35.2%; Pred. No. 3,6e-65;
 Matches 215; Conservative 100; Mismatches 246; Indels 49; Gaps 16;

QY 44 MDTHEPKLMNEDSYGNNEALLPMLMLAAGMOSYSORNSLDDIDIAENNDWILS 103
 DB 1 MESRPAKL-----YGAADARGPAVERLKTNLAAAHARPPPLALDAARADAPAMLHA 53
 QY 104 NKQVGVGVYDLPAGDLKGLKDKIPFQELGLTYLHMLPFCPEKSGSGGYAVSSYRDV 163
 DB 54 PQGTATFTYDVPFAGDINRGKLDYLTBLGYRWLHPLPLEPRGDSGSGRAVDYKY 113
 QY 164 NPALGTIGDLREVIAALHEAGISAVVDFIENHTSNEHEWAQCAAGDPLFDFNYIIFPD 223
 DB 114 DPLRTGIDLEALAGLRQDMGLILDVVCNHTAEHMAAARAGDPAYRDYITLPPDA 173
 QY 224 RMPDQYDRTRELFPDQHPGGS-QLIEDGRWYTTNSQOMLANSNPVFPAMGEMLF 282
 DB 174 QSAARARDELIVFPPTAGSFTYDAMAGGYWTFPFQMDLNPANPVFEMLEVLIF 233
 QY 283 LANLGVDLIRMDVAFIRKMGSTCENTLPQAHALIRAFNAVRIRIAPVFFKSEALVHPD 342
 DB 234 LIAKGAQGRLEDSAPFLMKQAGTTCNLDPTYEIVHAMAALSTIAPGVVLAELAESV 293
 QY 343 QVVOYIGODE--COIGYNPLQMALIMTLATREVNLLHQALTYRHNLPEHTAMVNVRS 400
 DB 294 DVLPPFGESGSCNLAANNVMTALMAALADGAVIARCLAVAAKPKQAGMLNVYRCH 353
 QY 401 DDTGTFADDDAAYLGISGYDRQFANRFVNRFPDGSFARGVPQYN---PSTGDCARS 456
 DB 354 DDLIMNAL---AAYAPAS---DIRWMSNAY--NGNGE--FSRGRAFOAEGGVST---N 400
 QY 457 GTAALVGLAQD---DPAVADRIKLIYSIALSTGGLPLTYLGDDEVTLANDDDWSODSNKS 513
 DB 401 GMAAALAGLTDADGDCGARRIRLTYGIIHALDGNPLTYMGDEIGLDNDAYQDDPLRA 460
 QY 514 DDSRWARPRYNEALYAQRNDPSTAGQTY-----QGLRHMIAVRQSNPRPDGRLVT 566
 DB 461 GGRMWHRPQDMSLAERGEAGALQADPFATFARLGGARLATLGVAGP---ARPV- 515
 QY 567 FNTNNKHIIIGYRNNNA---LIAFGNSEYPTVTAHTLQAMPKADHILGKTIVSLNQL 623
 DB 516 -EVSSEAVYAFIRDEGARBPFLCVANVSADAPQDF--ELPPAFAQAGEDVIDG--APSPAGAV 571
 QY 624 TLQGYVMWL 633
 DB 572 SLPPGIGTWL 581
 RESULT 9
 Q88FNO PRELIMINARY; PRT; 1106 AA.
 AC Q88FNO; 01-JUN-2003 (Tremblrel. 24, Created)
 DT 01-JUN-2003 (Tremblrel. 24, Last sequence update)
 DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
 DE Trehalose synthase, putative.
 GN PP4059.
 OS Pseudomonas putida (strain KT2440).
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;

OC Pseudomonadales; Pseudomonas.
 RN NCBI_TaxID=160488;
 RX SEQUENCE FROM N.A.

RA MEDLINE=22423060; PubMed=12534463;
 RA Nelson K.E., Weinel C., Paulsen I.T., Dodson R.J., Hilbert H.,
 RA Martins dos Santos V.A.P., Fouts D.E., Gill S.R., Pop M., Holmes M.,
 RA Brinkac L., Beanan M., DeBoy R.T., Daugherty S., Kolonay J.,
 RA Madupu R., Nelson W., White O., Peterson J., Khouri H., Hance I.,
 RA Chris Lee P., Holtzaple E., Scanlan D., Tran K., Moazed A.,
 RA Ueberback T., Rizzo M., Lee K., Kosack D., Moestl D., Wedler H.,
 RA Lauber J., Stjepandic D., Hohnesl U., Streetz M., Helm S.,
 RA Klewitz C., Eisen J., Timmis K.N., Dueserhoert A., Tuelmaier B.,
 RA Fraser C.M.;
 RT "Complete genome sequence and comparative analysis of the
 RT metabolically versatile *Pseudomonas putida* KT2440.";
 RL Environ. Microbiol. 4:799-808(2002).
 DR EMBL: AE016789; AAN69649.1; --
 DR TIGR: PP4059; --
 DR GO: GO:0004556; P:alpha-amylase activity; IEA.
 DR GO: GO:0005524; P:ATP binding; IEA.
 DR GO: GO:0003910; P:DNA ligase (ATP) activity; IEA.
 DR GO: GO:0005975; P:carbohydrate metabolism; IEA.
 DR GO: GO:0006310; P:DNA recombination; IEA.
 DR GO: GO:0006281; P:DNA repair; IEA.
 DR GO: GO:0006260; P:DNA replication; IEA.
 DR InterPro: IPR006047; Alpha.amyl.cat.
 DR InterPro: IPR000977; DNA ligase.
 DR Pfam: PF00128; alpha-amylase; 1.
 DR Prosite: PS00697; DNA_ligase_A1; 1.
 KW Complete proteome.
 KM SEQUENCE 1106 AA; 125589 MW; 2F5E471637096B38 CRC64;

Query Match 16.1%; Score 550; DB 16; Length 1106;
 Best Local Similarity 29.0%; Pred. No. 5,5e-35;
 Matches 164; Conservative 83; Mismatches 228; Indels 90; Gaps 20;

QY 118 GDLKELKDKIPFQELGLTYLHMLPFCPEKSGSGGYAVSSYRVNPLGTIGLREVI 177
 DB 39 GDFAGLISKLIDYIAELAGVNTLTLPEFY--PSPRRDGDYDIARVYKVAHPDYGSMADARRPI 96
 QY 178 AALHAGISAVVDFIENHTSNEHEWAQ---CAAGDPLFDFNYIIFPDGRMPDQ-YDRTL 233
 DB 97 AEHRKGRARVITELVINTSDHPFORAKRASKSAR--FYWMSD---DDQKYDGL- 150
 QY 224 REIFPDQHPGGS-QLIEDGRWYTTNSQOMLANSNPVFPAMGEMLFANLGVDLIR 292
 DB 151 RIIFIDTEKSMWMTQPVAGQYFMRHFGSHQDPLNDPNQVLKAVIGVRFMLDLGVDGLR 210
 QY 293 MDVAFIRKMGSTCENTLPQAHALIRAFNAVRIRIAPVFFKSEALVHPDQVVOYITGQ-- 350
 DB 211 IDAIFYLTERDGTNNENLAETHDIVKALRAEIDANYPPRMILAEANOWPEDTRFPYGBGD 270
 QY 351 -DECOIGYNPLQMALIMTLATREVNLLHQALTYRHNLPEHTAMVNVYRSHDQMTF-A 408
 DB 271 GDECMARHPPLMPMAYALAMEDFPITDILRCQPEIRANQMALFLRNDELLEWMT 330
 QY 409 DEDAAYLGISGYDRQFANRFVNRFPDGSFARGVPQYNPSTGDCRVS- GTAALVGLAQ 467
 DB 331 DRERDYL-----MNYAEDRRARINIGIRRLAPLQ 362
 QY 468 DDPHAVDRIKLIYSIALSTGGLPLTYLGDDEVTL-----LNDDD-----WSQDSN- 511
 DB 363 RDRR---RIELTSLILSPGTFTLYGDELGMGNITYGDRDGYRTPMQSPDRNGGS 419
 QY 512 KSDSRMARPRYNEALY-----AQRNDPSTAGQTYGGLRHMIAVRQSNPRPDGR 563
 DB 420 RADPQRLV-LPIMPPLVGYQVNVNVAQSHDPS---LNMTRRLTAVRKQKQARGGS 474
 QY 564 LVTFNNKHIIIGYR-----NMALIAFGNSEYPTVTAHTLQAMPKADHILGK 615
 DB 475 LRTLPNSRRILAIYREYTDAGNTEVILCVANVSRAAQAELLSQYADKVPVEMTLGGS 534

```

QY 616 TV-----SLNDLTLQPYQVWMLT 636
DB 535 APPPIGOLPFLTLPPYAFYWEFLA 559

RESULT 10
ID Q8XT75 PRELIMINARY; PRT; 1160 AA.
AC Q8XT75;
DT 01-MAR-2002 (TReMBLrel. 20, Created)
DT 01-MAR-2002 (TReMBLrel. 20, Last sequence update)
DT 01-JUN-2003 (TReMBLrel. 24, Last annotation update)
DE Hypothetical protein Rsp0240.
GN RSP0240 OR R505183.
OS Ralstonia solanacearum (Pseudomonas solanacearum).
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Burkholderiaceae; Ralstonia.
OK NCBI_Taxid=305;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=GM11000;
RX MEDLINE=21681879; PubMed=11823852;
RA Salanoubat M., Genin S., Artiguenave F., Gouzy J., Mangenot S.,
RA Arlet M., Billault A., Brottier P., Camus J.C., Catolico L.,
RA Chaudler M., Choisme N., Claudel-Renard C., Cumnac S., Demange N.,
RA Gaspin C., Lavye M., Moisan A., Robert C., Saurin W., Schlex T.,
RA Signier P., Tebault P., Whalen M., Winkler P., Levy M.,
RA Weisenbach J., Boucher C.A.;
RT "Genome sequence of the plant pathogen Ralstonia solanacearum.";
RL Nature 415:497-502 (2002).
DR EMBL; AL646077; CAD17391.1; -.
DR GO; GO:0046821; C:extrachromosomal DNA; IEA.
DR GO; GO:0004556; F:alpha-amyase activity; IEA.
DR GO; GO:0005975; P:carbohydrate metabolism; IEA.
DR InterPro; IPR006047; Alpha_amy1_cat.
DR InterPro; IPR006589; Alp_amy1_cat_sub.
DR Pfam; PF00128; alpha-amyase; 1.
DR SMART; SM00642; Amy; 1.
DR KMW plasmid; Hypothetical protein; Complete proteome.
SQ SEQUENCE 1160 AA; 129108 MW; 27555FA3ACEBDE5 CRC64;

Query Match 16.1%; Score 548; DB 16; Length 1160;
Best Local Similarity 29.3%; Pred. No. 8.5e-35;
Matches 164; Conservative 79; Mismatches 229; Indels 88; Gaps 20;

QY 118 GDLKHKDKIPYQOELTYHLMPFKCEGKSDGVAVSYVDNPAIGTIGDLREVI 177
DB 38 GDEPGISLKDYLAEIGVDAIWLPPY--PSPRDDGYDAEYRGVHPGYTMADARRFI 95

QY 178 AALHEAGISAVVDFINHTSNEHEW--ACRCAAGDPLFDNFYIIFPDRMPDQYRTLR 234
DB 96 AEAHAGLRVITELVINHTSDQHPFORARAKAGALRD--FYVMSD--DKXYAGR 150

QY 235 ELPFDQHPGFS--QLEDGRWVTTPNSPQMDLNYNPNWFRMAEMFLANLGVDIIR 252
DB 151 IIFIDEPSNMTWDPAADA--YVHRYSHQPDLPNPNRYLKAVLGVKFWMLNIGVDGLR 209

QY 223 MDVAFTWKOMGSCENTLPQAHALIRAFNAVMRIAPAVFEKSEALVHPDQVVQYGO-D 351
DB 210 LDVAFTVVEREGANENLPEHTAVLKIRAAAMAEKRNLLAKAQMWPEDQYENGND 269

QY 352 ECGIYNPLQWALLMNTLATREYNLLHQALTYYHNLPEHTAVNYVRSDDIGMTF-ADH 410
DB 270 ECHMAHFPLMPMYAIAREDRFPITDIRQTPPEVPPTCOMAIFLNHDELTLNEMVTA 329

QY 411 DAAYL-GISGYDHRQPLNRFVVRPDSFARGVFPQVNPSTGDCRVGTAALVGLAOD 469
DB 330 ERYLMEVVASDRRARN-----LGRRLRLALLRD 361

QY 470 PHAVDIRIKLYSTALSTGGLPLIYLGDEVGT-----LNDSD-----WSQDN---KSD 514
DB 362 RR---RVELMNSLGFMPGTPVWYVYDGEIOMGDNHILGPRDGVRTPMQNSPDINGFSRA 418

```

```

QY 515 DSRNAHRPRVNEALY-----AQRNDPSTAAQIYQGLRMHIAVROSNPRFDGRLVT 566
DB 419 DPEQLVPLAINGSLTYGSAVNEAQRD---ASHLWMTTRLLTRKRRHFRFGGSIOF 474

QY 567 FNTNKKHIIIGYR-----NNALAFGNFSEYQTV---TAHTLOAMPFKAHDLIGKTV-- 617
DB 475 LQPNARKYLAARALEGEAPILCVANLSRASQAVELDLAAVQORVP---ELIGTAFFP 531

QY 618 --SLNDLTLQPYQVWMLT 635
DB 532 IGLPYLTLPPYAFYWEFL 551

RESULT 11
ID Q911W3 PRELIMINARY; PRT; 1100 AA.
AC Q911W3;
DT 01-MAR-2001 (TReMBLrel. 16, Created)
DT 01-MAR-2001 (TReMBLrel. 16, Last sequence update)
DT 01-JUN-2003 (TReMBLrel. 24, Last annotation update)
DE Probable trehalose synthase.
GN PA2152.
OS Pseudomonas aeruginosa.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OK NCBI_Taxid=287;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 15692 / PAOI;
RX MEDLINE=20437337; PubMed=10984043;
RA Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warrenner P.,
RA Hickey M.U., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
RA Garber R.L., Goltzy L., Tolentino E., Westbrock-Wadman S., Yuan Y.,
RA Brody L.L., Coulter S.N., Folger K.R., Kass A., Laribi K., Lim R.M.,
RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
RA Reizer J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.;
RT "Complete genome sequence of Pseudomonas aeruginosa PAOI, an
RT opportunistic pathogen.";
RL Nature 406:959-964 (2000).
DR EMBL; AE004642; AAG05540.1; -.
DR PIR; G83376; G83376.
DR HSSP; P21332; IUOK.
DR GO; GO:0004556; F:alpha-amyase activity; IEA.
DR GO; GO:0005975; P:carbohydrate metabolism; IEA.
DR InterPro; IPR006047; Alpha_amy1_cat.
DR InterPro; IPR006589; Alp_amy1_cat_sub.
DR Pfam; PF00128; alpha-amyase; 1.
DR SMART; SM00642; Amy; 1.
DR KMW Complete proteome.
SQ SEQUENCE 1100 AA; 124521 MW; 00B2B8C8BDA8AF CRC64;

Query Match 15.9%; Score 541.5; DB 16; Length 1100;
Best Local Similarity 26.9%; Pred. No. 2.6e-34;
Matches 165; Conservative 90; Mismatches 223; Indels 135; Gaps 25;

QY 96 NNPDWILSKQVGVY---VDLF-----AGDLKDKIPYQOELTYHLMLDLK 145
DB 11 NNPQWTKD-----AVLYQVHVSFYDANNNGIGDFAGLEKLDYADLGVNTLWLPFY- 64

QY 146 CPEGKSDGVAVSYVDNPAIGTIGDLREVI AALHEAGISAVVDFINHTSNEHEW--- 202
DB 65 -PSPRDDGYDAIQAHYGVSIDYGLADARRFLAHRRLRVITELVINHTSDQHFWFR 123

QY 203 ACRCAAGDPLFDNFYIIFPDRMPDQYRTLRLEIPDQHPGFSQLEDGRWVT----- 256
DB 124 ARHAKKGBARD--YVMSD--SDEKYGST-RLIIFD-----TEGSNMTWDPAQOY 170

QY 257 ---TENSFQMDLNYSNPNWFRMAAGEMFLANLGVILMDVAFTWKOMGSCENTLPQ 313
DB 171 YVHRYSHQPDLPNPNRYLKAVLGVKFWMLNIGVDGLR-----LGRRLRLALLRD 361

QY 314 HALIRAFNAVMRIAPAVFEKSEALVHPDQVVQYIG-----QDECOIYNPLQWALLMN 367

```

Db 231 HOVLRIRAELEADLPDRMLAABANQMPEDTPTFYGGEGGEGGECCHNAFHPPLMPRTM 290
QY 368 TLATREVNLLHQALTYRHNLPHRTAMVNVSRSHDDIGWTFADEDAAYLGISGYDHRQPLN 427
Db 291 ALAQEDRYITDILQRTPTIPANCQWALFLRHDELTEMTVD-----DERDYL- 339
QY 428 RFFVNRFDGSEFARGVPFQYNPSTGDCRVS--GTAALVGLAODDPHAYDRIKLLYSAL- 484
Db 340 -----WNHYAADRRPALNLGIRRRRLPLVERRR--RIELHSLIL- 378
QY 485 SNGGFLTYLGDVEGT-----LNDD-----WSODSN-----KSDDSMARPRYNEA- 528
Db 379 SNGGFLTYLGDVEGT-----LNDD-----WSODSN-----KSDDSMARPRYNEA- 528
QY 529 Y-----AQRNDPSTAAGIYQGLRHMIAYRQSNPREDDGRVLTFTNNKHIIGYT-- 578
Db 438 YGQITINVEAQAQDPHS-----LNNMRRLLAVRSQOKKAGRSLKMLAPSNRIIAYLRE 493
QY 579 -----RNALLAFGNSEYPTQYTAHTLQAMPKXH-----DLIGKTV-----SLNDQ- 623
Db 494 YAEGRKOSILCVANLSRAAQA-----ELDLASHACKVPEMIGMSFPPIGELTYL- 547
QY 624 TLQPYQVMWLEIA 636
Db 548 TLQPYQVMWLEIA 560

RESULT 12

Q7W118 PRELIMINARY; PRT: 1113 AA.

Db 01-OCT-2003 (TREMBLrel. 25, Created)
Dt 01-OCT-2003 (TREMBLrel. 25, Last sequence update)
Dt 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Alpha-glucosidase.
GN BB2863.
OS Bordetella bronchiseptica (Alcaligenes bronchiseptica).
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Alcaligenaceae; Bordetella.
OX NCBI Taxid=518;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=RB50 / ATCC BAA-588;
RX MEDLINE=22827954; PubMed=12910271;
RA Parthill U., Sebathia M., Preston A., Murphy L.D., Thomson N.,
RA Harris D.E., Holden M.T.G., Churcher C.M., Bentley S.D., Mungall K.L.,
RA Cerdano-Tarraga A.M., Temple L., James K., Harris B., Quail M.A.,
RA Achman M., Atkin R., Baker S., Basham D., Bason N., Chervach I.,
RA Chillingworth T., Collins M., Cronin A., Davis P., Doggett J.,
RA Fellwell T., Goble A., Hamlin N., Hauser H., Holroyd S., Jagels K.,
RA Leather S., Moule S., Norbertczak H., O'Neill S., Ormond D., Price C.,
RA Rabinowitsch E., Rutter S., Sanders M., Saunders D., Seeger K.,
RA Sharp S., Simmonds M., Skelton J., Squares R., Squares S., Stevens K.,
RA Unwin L., Whitehead S., Barrell B.G., Maekell D.J.,
RT "Comparative analysis of the genome sequences of Bordetella pertussis,
RT Bordetella parapertussis and Bordetella bronchiseptica."
RL Nat. Genet. 35:32-40(2003).
DR EMBL; BX640445; CAE33355.1; --
KW Complete proteome.
SQ SEQUENCE 1113 AA; 124163 MW; APCS2AB054940D1 CRC64;

Query Match 15.9%; Score 541.5; DB 16; Length 1113;

Best Local Similarity 27.2%; Pred. No. 2.7e-34; Indels 99; Gaps 17;

Matches 154; Conservative 90; Mismatches 223;

Db 118 GDLKGLKDKIPYFQELGLTYLHMLPFCPCBEKSGDGYAVSSYRDVNPALGTIGDLREVI 177
QY 35 GDLKGLKDKIPYFQELGLTYLHMLPFCPCBEKSGDGYAVSSYRDVNPALGTIGDLREVI 92
Db 178 AALHENGASAVVDYFNHTSNHEWMAQRCAGDP--LFDNFYIIPDRAMPDQYDRTLEI 236
QY 93 RAAHARGALVTELVINHTSDQHPPWFQBARAARPGSAHRAVYVMSD---DKAVAGTIRII 149

QY 237 PPDQHPGFS-QLEDGRVWTTTNSFQMDLAYSNPWVRMAAGMLFIANLCYDILRMDA 295
Db 150 FCDERKSNMTDVPAGAFEMRFYSHQDPLNDYNDQYREVLAGARRYMLDNGVDRLRDA 209
QY 296 VAFIKWKGSTCENLPQAHALIRAFNAVMRIAAPAVFKESEAIYHPDVOVYIQG-DECO 354
Db 210 VPIYVERGNTNNENPEHTAILRRIRRYIDSEYFPRMLAABANQMPEDAQAETFGAGDECH 269
QY 355 IGYNEPLQALLMNTLATREVNLLHQALTYRHNLPHRTAMVNVSRSHDDIG----- 404
Db 270 MAHFELPLMRMYMALAQSDRLPYVDII IQGTSIAQCQWALFLRHDELTEMTVSREND 329
QY 405 --W-TFADEDAAYLGISGYDHRQPLNRFVNRFDGSEFARGVPFQYNPSTGDCRVSATAA 461
Db 330 YLWNYAAEPFARINL-----GIRRR 350
QY 462 LVGLAODDPHAYDRIKLLYSALSTGGLPLTYLGDVEGT-----LNDD-----WSOD 509
Db 351 LAPLERDR--RIELNSLILSMPTPVLYGDELMGDNHILGDRDGVATTPMQWSPD 407
QY 510 SN-----KSDDSMARPRYNEALY-----AQRNDPSTAAGIYQGLRHMIAYRQSNP 557
Db 408 RNGGFSRADPBR--LEPLMGLPLGYEAVNVEAQAQDPHS-----LNNMTRMLARROSH 462
QY 558 RPDGRLVTFNTNNKHIIGYT--NNALLAFGNSEYPTQYTAHTLQAMPKXH----- 614
Db 463 VFGRELSFLYPGNKILAYRTMEDIVLVANLSQAQAQVEHLSEYAGVPEMIGMS 522
QY 615 KTV-----SLNDQTLQPYQVMWLEIA 636
Db 523 TAPQIGLPLYLITPFGFWLDS 548

RESULT 13

Q7VYJ9 PRELIMINARY; PRT: 1113 AA.

Db 01-OCT-2003 (TREMBLrel. 25, Created)
Dt 01-OCT-2003 (TREMBLrel. 25, Last sequence update)
Dt 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Alpha-glucosidase.
GN BP1329.
OS Bordetella pertussis.
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Alcaligenaceae; Bordetella.
OX NCBI Taxid=520;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Tohama I / ATCC BAA-589 / MCTC 13251;
RX MEDLINE=22827954; PubMed=12910271;
RA Parthill U., Sebathia M., Preston A., Murphy L.D., Thomson N.,
RA Harris D.E., Holden M.T.G., Churcher C.M., Bentley S.D., Mungall K.L.,
RA Cerdano-Tarraga A.M., Temple L., James K., Harris B., Quail M.A.,
RA Achman M., Atkin R., Baker S., Basham D., Bason N., Chervach I.,
RA Chillingworth T., Collins M., Cronin A., Davis P., Doggett J.,
RA Fellwell T., Goble A., Hamlin N., Hauser H., Holroyd S., Jagels K.,
RA Leather S., Moule S., Norbertczak H., O'Neill S., Ormond D., Price C.,
RA Rabinowitsch E., Rutter S., Sanders M., Saunders D., Seeger K.,
RA Sharp S., Simmonds M., Skelton J., Squares R., Squares S., Stevens K.,
RA Unwin L., Whitehead S., Barrell B.G., Maekell D.J.,
RT "Comparative analysis of the genome sequences of Bordetella pertussis,
RT Bordetella parapertussis and Bordetella bronchiseptica."
RL Nat. Genet. 35:32-40(2003).
DR EMBL; BX640415; CAE41622.1; --
KW Complete proteome.
SQ SEQUENCE 1113 AA; 124135 MW; 7235BDC938E49756 CRC64;

Query Match 15.8%; Score 538.5; DB 16; Length 1113;

Best Local Similarity 27.2%; Pred. No. 4.7e-34; Indels 99; Gaps 17;

Matches 154; Conservative 90; Mismatches 223;

Db 118 GDLKGLKDKIPYFQELGLTYLHMLPFCPCBEKSGDGYAVSSYRDVNPALGTIGDLREVI 177

```

Db 35 GDFAGLAKLDYIVLGVNTTWLFFY--PSPRRDDGDIADYKRVHVDYGSIAADALIV 92
Qy 178 AALHAGISAVVDFLEFNHSHNEHWAQRCACADP-LFDNFYIIFPDRMPQDYETLEI 236
Db 93 RAARHARGRVITELVINTSDQHPFORARARPCSAHRAAYVMSD---DDKAYAGTRLI 149
Qy 237 FPDQHPGGS-OLEGKRWVWTFNFSFQNDLNTSNWVRAMAGELFLANLGVDTLIRDA 295
Db 150 FCDDEKSWTWDPVAGAFYFMRFYSHQDPLNDNQVLEVIYGAARWYLDKGVDELRLDA 209
Qy 296 VAFIWKQWTSCEMPOAHALIRAFNAVMRIAAVAFKSEAIYHPDQVQYIQ-DECO 354
Db 210 VPIYVERGNTNENIPETHALIRKRVYIDSEYPERMLAEBANQWPERAQBYFGAGDECH 269
Qy 355 IGYNPLQWALIMTLATREVNLMHQAITYRNHLEPHITAMVNVYRSHDDIG----- 404
Db 270 MAHFHPLMRMYMALAQEDRLPVTDIIRQTPSIAPQCOWALFLRNHDELILEMTYRBRD 329
Qy 405 --W-TFAEDDAAYIGISGYDRQPLNRFVNRFDGSPARGVFGQYNPSTGCRVSGTAA 461
Db 330 YLMWVYAEPPARIML-----GIRRR 350
Qy 462 LVGLAODDPHAVDRIKLYSIALSTGGLPLIYIGDEVGT-----LNDD-----MSQD 509
Db 351 LAPLEFRRR--RIELMNSLILSPGTPLYIGDELKGNINHLGDRDQVRTFMQNSPD 407
Qy 510 SN---KSDSRMAHRRPYNEALY-----AQNDPSTAAGQIYQGLRHMAVQSNP 557
Db 408 RNGGFSRADPER-LEPLFMGLPLYEAVNEAQAQRDPS-----LIMNTRMLAKRRQSH 462
Qy 558 RPDGRLVTFNTNNHIIIGYR--NNALIFAGNSEFPQVYATLTLOMFPKADLIG 614
Db 463 VFGRELSEFLYPGNKILAYLRTWEDTVLGVANLSQAAPVELLSYAGRVPEMLIG 522
Qy 615 KTV---SLNDLTLQPYQVMMLBIA 636
Db 523 TAPFOIGLPIYLLIPPGFTYLDLS 548

RESULT 14
054207 PRELIMINARY; PRT; 572 AA.
AC 054207,
DT 01-JUN-1998 (TRENBLrel. 06, Created)
DT 01-JUN-1998 (TRENBLrel. 06, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE Putative trehalose synthase.
GN SC07334 OR SC04G10.13C.
OS Streptomyces coelicolor.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycinae; Streptomycetaceae; Streptomycetes.
OX NCBI_TaxID=1902;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RA Schneider D., Bruton C.J., Chater K.F.;
RL Submitted (MAR-1998) to the EMBL/Genbank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RA MEDLINE=96154943; PubMed=8596463;
RX Bruton C.J., Plakitt K.A., Chater K.F.;
RT "Tissue-specific glycosyl branching isoenzymes in a multicellular
RL Mol. Microbiol. 18:89-99 (1995).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RA Seeger K.J., Harris D.;
RL Submitted (MAY-2000) to the EMBL/Genbank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.

```

```

RC STRAIN=A3(2);
RA Thomson N.R., Parkhill J., Barrell B.G., Rajandream M.A.;
RL Submitted (MAY-2000) to the EMBL/Genbank/DBJ databases.
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RX MEDLINE=97000351; PubMed=8843436;
RA Redenbach M., Kleiser H.M., Denapate D., Eichner A., Cullum J.,
RA Kinash H., Hopwood D.A.;
RT "A set of ordered cosmid and a detailed genetic and physical map for
RL the 8 Mb Streptomyces coelicolor A3(2) chromosome."
RN [6]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2) / M145;
RX MEDLINE=2196410; PubMed=12000953;
RA Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,
RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kleiser H.,
RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
RA Huang C.-H., Kleiser T., Larke L., Murphy L., Oliver K., O'Neill S.,
RA Rabinowitsch E., Rajandream M.A., Rutherford K., Rutter S.,
RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
RA Warren T., Wietzorrek A., Woodward J., Barrell B.G., Parkhill J.,
RA Hopwood D.A.;
RT "Complete genome sequence of the model actinomycete Streptomyces
RL coelicolor A3(2).";
RL Nature 417:141-147 (2002).
DR EMBL; AJ001206; CAA04607.1; -
DR EMBL; AL939131; CAB92880.1; -
DR HSSP; P21332; 100K.
DR GO; GO:0004556; P:alpha-amylose activity; IEA.
DR GO; GO:0004553; F:hydrolase activity; hydrolyzing O-glycosyl . . . IEA.
DR GO; GO:0005975; P:carbohydrate metabolism; IEA.
DR InterPro; IPR006047; Alpha_amy1_cat.
DR InterPro; IPR006589; Alp_amy1_cat_sub.
DR InterPro; IPR008263; Glyco_hyd16_AS.
DR Pfam; PF00128; alpha-amylose; 1.
DR SMART; SM00642; Aamy; 1.
DR PROSITE; PS01034; GLYCOSYL_HYDROL_F16; 1.
DR Complete proteome.
SQ SEQUENCE 572 AA; 65695 MM; 81252A982BBD51D CRC64;

Query Match 15.5%; Score 529.5; DB 16; Length 572;
Best Local Similarity 26.9%; Pred. No. 8.8e-34;
Matches 161; Conservative 85; Mismatches 244; Indels 109; Gaps 19;

Qy 93 ARENNPDI-----LSNKQYGVYCYVDLAFAGDKLKGDKXIPYFQELGTYLTL 140
Db 16 AGDHPDFKRAVYEVLYVRSFQDSNGDI-----GDLAGLTKADYLOMIGVDCIML 68
Qy 141 MFLKCPGKSDGYAVSSYRDVPALGTIGDLREVIATLHAGISAVVDFLEFNHSHNEH 200
Db 69 PPFKSP--LMDGVDVSDYAVLPEFGDLADFVEFVDAHQRMVRIIDVNMHTSDQH 126
Qy 201 EMAQ--RCAAGDPLFDNFYIIFPDRMPDQYDRLRLRFPPQHGGGSEQLD---GRVW 255
Db 127 PWFQESKRPDGPAGDYVWADDDTRYAD-----ARILFVTEASNNW--HDPVARGQYW 179
Qy 256 TTFNSFQMDLNTSNWVFRAMAGEMFLANLGVDTLIRDAVAFIWKQWTSCEMPOAHA 315
Db 180 HRFESHQDPLNTSNWVFRAMAGEMFLANLGVDTLIRDAVAFIWKQWTSCEMPOAHA 239
Qy 316 LIRAFNAVMRIAAVAFKSEAIYHPDQVQYIQ-----DECOIGYNPLQWALIMTLA 370
Db 240 FLKVRREIDRQYDPTVLAEBANQWPERVVDVYFGDYSTGDECHMAFFPVMPRIEMAVR 299
Qy 371 TREVNLMHQAITYRNHLEPHITAMVNVYRSHDDIGTF-ADDDAAYLGISGYDRQPLNRF 429
Db 300 RESRYPVSEILAKTPALPSCQWGIPLRNHDELILEMTYDERDYM----- 345
Qy 430 FVNRFDGSPARGVFGQYNPSTGCRVS---GTAALVGLAODDPHAVDRIKLYSIALST 486

```

```

Db 346 -----YAEVAKDPRMRANIGIRRRRLATLNDNR---DQIEFTALLAL 386
QY 487 GGLPLIYLGEVGT-----LNDND-----WSQDSN---KSDSRMAHRRPVNEALY-- 529
Db 387 PGSILIVYGDIGMGNDINWLGDRDAVRTPMQMTDPRNAGFSTCDPRLYLPAIMDPYGY 446
QY 530 -----AQRNDPSTAAGIYQGLRHMTAVROSNEPFDGRLVTNTNKKHIIGYR---N 580
Db 447 QVTNVEASMSPPSS-----LHWTBRMIEIRKQNPAGFGLGYTELPSNPAYLAFLREYED 502
QY 581 NALLAFGNFSEYPTQVTAHTIQAMPFKHDLIGKTV-----SLNQDLTLQYQVWMLEI 635
Db 503 DLVLCVNNFARFAQPTLDELREFAGRHPLELFGVRFPAIGELPYLLTLGGHGFYFRL 561

RESULT 15
082JF2 PRELIMINARY; PRT; 572 AA.
ID 082JF2 AC 082JF2;
DT 01-JUN-2003 (TREMblrel. 24, Created)
DT 01-JUN-2003 (TREMblrel. 24, Last sequence update)
DT 01-OCT-2003 (TREMblrel. 25, Last annotation update)
DE Putative trehalose synthase.
GN SAV2803.
OS Streptomyces avermitilis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycinae; Streptomycetaceae; Streptomyces.
CX NCBI_TaxId=33903;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;
RX MEDLINE=21477403; PubMed=11572948;
RA Omura S., Ikeda H., Ishikawa J., Hanamoto A., Takahashi C.,
RA Shinose M., Takahashi Y., Horikawa H., Nakazawa H., Osone T.,
RA Kikuchi H., Shiba T., Sakaki Y., Hattori M.;
RT "Genome sequence of an industrial microorganism Streptomyces
RT avermitilis: deducing the ability of producing secondary
RT metabolites.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220 (2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;
RX MEDLINE=22608306; PubMed=12692562;
RA Ikeda H., Ishikawa J., Hanamoto A., Shinose M., Kikuchi H., Shiba T.,
RA Sakaki Y., Hattori M., Omura S.;
RT "Complete genome sequence and comparative analysis of the industrial
RT microorganism Streptomyces avermitilis.";
RL Nat. Biotechnol. 21:526-531 (2003).
DR EMBL, AP005032; BAC70514.1;-.
DR GO; GO:0004556; F:alpha-amyase activity; IEA.
DR GO; GO:0005975; P:carbohydrate metabolism; IEA.
DR InterPro; IPR006047; Alpha_aml_cat.
DR InterPro; IPR006589; Alp_aml_cat_sub.
DR pfam; PF00128; alpha-amyase; 1.
DR SMART; SM00642; Amy; 1.
KW Complete proteome.
SQ SEQUENCE 572 AA; 65751 MW; A6B0EAB49FF4D3EA CRC64;

Query Match 15.5%; Score 529.5; DB 16; Length 572;
Best Local Similarity 27.2%; Pred. No. 8,8e-34;
Matches 166; Conservative 86; Mismatches 233; Indels 125; Gaps 23;

QY 93 ARENNPDWT-----LSNKQGVGYCYVDLPAGDLKGLKDKIPYQELGLTYLHL 140
Db 16 AKDRDPEWFKRAVFEVLRVSFQDSNGGV-----GDLKGLTANKLDYLTQWLGVDCLMT 68
QY 141 MLEKCPKESKSGYAVSYRDNVPALGTIGDLREVIALHRAGISAVVDFFIENHNSNEH 200
Db 69 PPEFFKSP--LRDGGYDVSDTYAVLPFGDLADFEVDAHQGRKRVIIIDFVMAHISDLH 126
QY 201 EWAQ--RCAAGDPLFDNFYIIPDRRMPQDYDRTLRKIFPDQHPGFSQLEBGRV--- 254
Db 127 PWFQESRSNPDSGPDYD--YYWAD---DDKQYQDARIIFVD-----TEASNWTDFPV 173

```

```

QY 255 -----WTENSFGMDLINSNPNWFRAMAGMELPLANIGVDILRNDAYAFIMKMGSTCEN 309
Db 174 RKQYWMRRFFPSHOPDINTYENPAVOEELVSALRFWLDIGIGFRIDAVPYLYQEGTNCEN 233
QY 310 LPQAHALIRAFNWMVRLAAVAFKFSRAIYHPQVQYIQQ-----DECOIGVPLQMAL 364
Db 234 LPATHEFLKVRKEIDTHYPTVLLAEANQMPEDVDVYFGDFSGGDECHMAFFFPVMPR 293
QY 365 LMTLATREYNLLHQALTYRHNLPEHTAWNYVYSHDDIGWTF-ADEDAAYLGISYDHR 423
Db 294 IFPAVRESRYPPSEILAKTPAIPSSCGMFLRNHDELTEWTDERDYMM----- 346
QY 424 QFLNRFVNRPDGSFARGVFPQYNPSTGCRVSGTAALVGLAODDPHADRILKLSIA 483
Db 347 -----AEYAKDPRMRAN-----IGIRRRRLATLNDNR---NQIEFTALL 383
QY 484 LSTGGLPLIYLGEVGT-----LNDND-----WSQDSN---KSDSRMAHRRPVNEAL 528
Db 384 LSLPGSPILYIGDELIGMGNDINWLGDRDAVRTPMQMTDPRNAGFSSCDPRLYLPTIMDPV 443
QY 529 Y-----AQRNDPSTAAGIYQGLRHMTAVROSNEPFDGRLVTNTNKKHIIGYR- 579
Db 444 YGQVTVNVEASMSPPSS-----LHWTBRMIEIRKQNPAGFGLGYTELQSNPAYLAFLRE 499
QY 580 -----NALLAFGNFSEYPTQVTAHTIQAMPFKH--DLIGKTV-----SLNQDLTL 625
Db 500 APSTGGNGDDLVLCVNNFARFAQPTLDELREFAGRHPLELFGVRFPAIGELPYLLTL 557
QY 626 QPYQVWMLEI 635
Db 558 AGHGFYFRL 567

```

Search completed: May 25, 2004, 19:31:55
Job time : 49 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 25, 2004, 18:56:26 ; Search time 17 Seconds

(without alignment)
1948.036 Million cell updates/sec

Title: US-09-807-146-1

Sequence: 1 MLTPQVGLILQYLKTRIL.....VSLNODLTLPQVQVWLEIA 636

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database: SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	504	14.8	573	1	TRES_P1MSR
2	484	14.2	563	1	TRES_THET
3	452.5	13.3	559	1	SUCP_ECOLI
4	375	11.0	586	1	MAL1_DROVI
5	330.5	9.7	524	1	MAL2_DROVI
6	325.5	9.6	567	1	MAL1_APIME
7	324.5	9.5	441	1	MGT1_THEMA
8	314	9.2	562	1	OL6G_BACTR
9	308	9.0	442	1	MGT1_THEMA
10	304	8.9	498	1	AMY3_DICTH
11	304	8.9	558	1	OL6G_BACCE
12	294	8.6	508	1	MAL1_DROME
13	292	8.6	520	1	AMY_BACME
14	289.5	8.5	555	1	OL6G_BACCO
15	289.5	8.5	537	1	DEXB_STREB
16	288.5	8.5	551	1	AGLA_RHIME
17	285	8.4	522	1	MAL2_DROME
18	281	8.2	505	1	MAL3_DROME
19	281	8.2	584	1	MAL3_YEAST
20	280	8.2	584	1	MAL3_YEAST
21	280	8.2	584	1	MAL3_YEAST
22	276.5	8.1	551	1	TREC_ECOLI
23	272	8.0	561	1	OL6G_BACCH
24	267.5	7.8	561	1	OL6G_BACCH
25	265	7.8	583	1	SC31_RAT
26	263	7.7	589	1	MAYS_YEAST
27	263	7.7	1279	1	APU_THESA
28	262.5	7.7	685	1	SC31_HUMAN
29	261.5	7.7	508	1	OL6G_BACSP
30	261	7.7	579	1	MALT_ABDAC
31	258.5	7.6	557	1	AGL_FEDPE
32	256.5	7.5	1475	1	APU_THET
33	253.5	7.4	1481	1	APU_THET

34	250	7.3	585	1	NEP2_THETV	006751 thermactin
35	247	7.2	589	1	MAL2_YEAST	P53051 saccharomy
36	245.5	7.2	535	1	DEXB_STREB	O54796 streptococ
37	242	7.1	536	1	DEXB_STREB	O09040 streptococ
38	242	7.1	569	1	MALT_CANAL	O02751 candida alb
39	240.5	7.1	561	1	TREC_BACSV	P39795 bacillus su
40	232	6.8	591	1	CDAS_BACSH	O08341 bacillus sp
41	231	6.8	619	1	AMY_STRLI	O05884 streptomyc
42	230	6.7	605	1	MAL2_ECOLI	P21517 escherichia
43	224.5	6.6	581	1	MAYS_YEAST	P40884 saccharomy
44	222.5	6.5	588	1	NEPU_BACST	P38940 bacillus st
45	222	6.5	1861	1	APU_THETV	P38536 t amylopull

ALIGNMENTS

RESULT 1	TRES_P1MSR	STANDARD	PRT	573 AA.
ID	TRES_P1MSR			
AC	P72235			
DT	28-FEB-2003 (Rel. 41, Last sequence update)			
DT	28-FEB-2003 (Rel. 41, Last sequence update)			
DT	10-OCT-2003 (Rel. 42, Last annotation update)			
DE	Trehalose synthase (EC 5.4.99.16) (Maltose alpha-D-glucosyltransferase).			
GN	TRES.			
OS	Pimelobacter sp. (strain R48).			
OC	Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;			
OC	Propionibacteriaceae; Nocardioidaceae; Pimelobacter.			
OK	NCBI_Taxid-51662;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE-96239784; PubMed-8645698;			
RA	Tsutsuki K., Nishimoto T., Nakada T., Kubota M., Chaen H., Sugimoto T.,			
RT	Kurimoto M.;			
RT	Cloning and sequencing of trehalose synthase gene from Pimelobacter			
RL	sp. R48."			
RL	Biochim. Biophys. Acta 1290:1-3 (1996).			
CC	-1- FUNCTION: Catalyzes the conversion of maltose into alpha, alpha-trehalose by transglucosylation.			
CC	-1- CATALYTIC ACTIVITY: Maltose = alpha, alpha-trehalose.			
CC	-1- SIMILARITY: Belongs to family 13 of glycosyl hydrolases.			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).			
CC	or send an email to license@isb-sib.ch .			
DR	EMBL; D78198; BAAL1303.1; -.			
DR	PIR; S71450; S71450.			
DR	HSSP; P21332; IUOK.			
DR	InterPro; IPR006589; Alp_amy1_cat_sub.			
DR	InterPro; IPR006047; Alpha_amy1_cat.			
DR	Pfam; PF00128; alpha-amy1ase; 1.			
DR	SMART; SM00642; Amy; 1.			
KW	Isomerase.			
SC	SEQUENCE 573 AA; 65668 MW; 904990XAA97AIE79 CRC64;			
QY	Query Match	14.8%; Score 504; DB 1; Length 573;		
QY	Best Local Similarity	25.6%; Pred. No. 3.9e-31;		
QY	Matches 155; Conservative 96; Mismatches 232; Indels 122; Gaps 23;			
Db	98 PMW	-----LSMKVGVGVYVDFPGDLKGLKDKIPYFQELGLTYHLMLPKR 145		
Db	13 PFMPRTAVFVYVYVMSFRPNAGG-----TGDFRGLKEDYDYLQWLVDCIMVPPFS 65			
QY	146 CPEGKSDGVAVSSRYDVNPALGTIGDLREVTALAEAGISAVVDPIFNHTSNEHW--A 203			
Db	66 SF--LRDGGYDVADVTGILPEIGYEDFPAFLDGAHERGIRVILIDVNMHTSDAHWFOA 123			


```

RA Blatner F.R., Plunkett G., Iii, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-VIDES J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12.";
RL [12]
RN accession 277:1453-1474(1997).
RP SEQUENCE FROM N.A.
RC STRAIN-K12;
EX MEDLINE=97251357; Pubmed=9097039;
FA Alpha H., Baba T., Fujita K., Hayashi K., Inada T., Isono K.,
RA Itoh T., Kasai H., Kashimoto K., Kimura S., Kitakawa M.,
RA Kitagawa K., Makino K., Miki T., Mizobuchi K., Mori H., Mori T.,
RA Momomura K., Nakabe S., Nakamura Y., Nashimoto H., Nishio Y.,
RA Oshima T., Saito N., Sampei G., Seki Y., Siyasundaram S.,
RA Tagami H., Takeda U., Takemoto K., Takeuchi Y., Yamada C.,
RA Yamamoto Y., Horinouchi T.;
RT "A 570-kb DNA sequence of the Escherichia coli K-12 genome
corresponding to the 28.0-40.1 min region on the linkage map.";
RL DNA Ref. 3:363-377(1996).
CC -1- CATALYTIC ACTIVITY: Sucrose + phosphate = D-fructose + alpha-D-
CC glucose 1-phosphate.
CC -1- SIMILARITY: Belongs to the sucrose phosphorylase family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC CC
CC EMBL; AE000229; AAC74391.1; ALT_INIT.
CC DR EMBL; D90768; BAAL4878.1; -.
CC DR EMBL; D90769; BAAL4886.1; -.
CC DR Ecocore; EGJ31910.ycJM.
CC DR InterPro; IPRO06047; Alpha_amy1_cat.
CC DR Pfam; PF00128; alpha-amylase; 1.
CC DR Hypothetical protein; Transferase; Glycosyltransferase;
CC KW Complete proteome.
CC SQ SEQUENCE 559 AA; 64185 MW; C05B71B0E0369C2 CRC64;

Query Match 13.3%; Score 452.5; DB 1: Length 559;
Best Local Similarity 26.7%; Pred. No. 3.5e-27;
Matches 165; Conservative 98; Mismatches 260; Indels 95; Gaps 26

QY KLAMELDVYGNNEALLPMLFMLLAQAQSORSNSKIDIDIAENNPWILSNKYGG 109
DB 4 KITTYLDIBIYGT-----FPAATHQLVTRBSAKPLITGRKKHW--DESDVVL 51
QY 110 VCAYVDLFAGDLKGDKIPYOEIG-----LTYHLMPFKCPBGSKDGVAVSSYR 161
DB 52 ITYADQFISNDLKPE---PTFNDFYHOMLOSIFSHVLLPFY---PWSSDDGFSVIDYH 104
QY 162 DVNPALEGTGLREVIALHAEGISAVVDFFLNFTSNHEWAQRCAADPLFDNYYYTFP 221
DB 105 QVASAEAGMWOLIQQLGECSH-----LMEDFYCNHMSASSEFPKNTLOQHPEDEFDT--- 156
QY 222 DRAMPDOYD-----RTUREIPDPDHQGFSGLED--GRVVWTTSFQMDLYNSFW 272
DB 157 --AADPDQDISAVTRPRAPLPILTP-----FQMKDSHTRLMTFTSDDQIDLNRSPEV 207
QY 273 FRAAAGEMFLPLANTGYDLIRMDVAVFIMKMGTSCEMTPQAHAIIRANAWRLTAAPVF 332
DB 208 LLAAWDVALCTIAGAETVRDAVGFMKEPGSTIHLEKTHIIKLRSITIINDVAPETV 267
QY 333 FKSEA-IVHPDVVOY-IGODECOIGNPLDMALLMNTLATREVNLT---HQALTFRNL 387
DB 268 IITETNVPHKNIVAFVGAGDEAHMVVOFSPLVIHAVQONVEALCAMQNLT---L 323
QY 388 P-EHTAWNVYRSDDIGMT-----FADEDAAVYGISGYDRKQPLNRFVNFPGSPARG 441
DB 324 PSSVTWNFNFLASHDGIGLNDLRGLPESELIEVLAQQOQGALVN--WKNNPDGTRS-- 379

```

Oy		442	VPEFYNSTGCGCRVSGTAALVLGLADDPNAVHDIKLVLSIALSTGGLPLILYLDGEVTL	501
Dd		380	-PYEINATYMD-----ALSRESSDEERCAFIHAALLSFPGPVALITYSIISR	430
Oy		502	NDDMOSDSDNSKDSRWAHPRPY-NBALYAQRNDPSTAAGIYGGLRHMIAVROSNPREF	560
Dd		431	ND---YAGVELTGYNRAINKKHYSKIITREINDAETLRHAVYHELSLITLRSHNEFH	487
Oy		561	GGRLVTEFTNNKHIIIGYIRNNA----LIAFNSEBYPQTVAHTLQAMFFAAHDLGSKT	616
Dd		488	PDNNTFTIDTIINSYMARIPRSNADGNCLTGLFNVSKNIOHVNIITMLH-----GRDLISEVD	542
Oy		617	VSUNODLTLPQRYVMWLE	634
Dd		543	I-LGNEITLRPMQVMWK	559
 RESULT 4 MAL1 DPROVI STANDARD; PRT; 586 AA.				
ID	MAL1 DPROVI	STANDARD;	PRT;	586 AA.
AC	O16096;			
DT	30-MAY-2000 (Rel. 39, Created)			
DT	30-MAY-2000 (Rel. 39, Last sequence update)			
DT	10-OCT-2003 (Rel. 42, Last annotation update)			
DE	Maltase 1 precursor (EC 3.2.1.20).			
OS	MALT1.			
CS	Drosophila virilis (Fruit fly).			
OC	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;			
OC	Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;			
OC	Ephydroidea; Drosophilidae; Drosophila.			
OX	NCBI_TaxID=7244;			
RN	[1]			
RC	SEQUENCE FROM N.A.			
RF	STRAIN=9;			
FX	MEDLINE=97475684; PubMed=9335139;			
RA	Vieira C.P., Vieira J., Hartl D.L.;			
RT	"The evolution of small gene clusters: evidence for an independent origin of the maltase gene cluster in Drosophila virilis and Drosophila melanogaster."			
RL	Mol. Biol. Evol. 14:985-993(1997).			
CC	-1- CATALYTIC ACTIVITY: Hydrolysis of terminal, non-reducing 1,4-linked D-glucose residues with release of D-glucose.			
CC	-1- SIMILARITY: Belongs to family 13 of glycosyl hydrolases.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.ebi.ch/announce/or send an email to license@ebi.ac.uk).			
CC	-----			
CC	EMBL; AF006573; AAB82327.1; -.			
DR	HSPB; P21332; IJOK.			
DR	FlyBase; FBgn0022839; Dvir\May1.			
DR	InterPro; IPRO06589; Alp_amy1_cat_sub.			
DR	InterPro; IPRO06047; Alpha_amy1_cat.			
DR	Pfam; PF00128; alpha-amylase; I.			
DR	SMART; Sm00642; Amy; I.			
KM	Hydrolase; Glycosidase; Signal.			
FT	SIGNAL	1	19	POTENTIAL.
FT	CHAIN	20	586	MALTASE 1.
FT	CARBOHYD	133	133	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	166	166	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	287	287	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	415	415	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	529	529	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	532	532	N-LINKED (GLCNAC. . .) (POTENTIAL).
SO	SEQUENCE	586 AA;	67544 MW;	FD88C625FAFC868 CRC64;

Query Match 11.0%; Score 375; DB 1; Length 586;

Best Local Similarity 23.1%; Pred. No. 3.ee-21;

OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Hymenoptera; Apocrita; Aculeata; Apoidea;
 OC Apidae; Apis.
 CX NCBI_TaxID=7460;
 RN [1]
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 18-41; 116-137 AND 255-267.
 RC TISSUE=hypopharyngeal gland;
 RX MEDLINE=96192057; PubMed=8619664;
 RA Ohashi K., Sawata M., Takeuchi H., Natori S., Kubo T.;
 RT "Molecular cloning of cDNA and analysis of expression of the gene for
 alpha-glucosidase from the hypopharyngeal gland of the honeybee *Apis
 mellifera* L.";
 RL Biochem. Biophys. Res. Commun. 221:380-385(1996).
 RN [2]
 RP CHARACTERIZATION.
 RC TISSUE=hypopharyngeal gland;
 RX MEDLINE=97037073; PubMed=8882720;
 RA Kubo T., Sasaki M., Nakamura J., Sasagawa H., Ohashi K., Takeuchi H.,
 RN Natori S.;
 RT "Change in the expression of hypopharyngeal-gland proteins of the
 worker honeybees (*Apis mellifera* L.) with age and/or role.";
 RL J. Biochem. 119:291-295(1996).
 CC -1- FUNCTION: Converts sucrose in nectar to glucose and fructose.
 CC -1- CATALYTIC ACTIVITY: Hydrolysis of terminal, non-reducing 1,4-
 linked D-glucose residues with release of D-glucose.
 CC -1- SUBUNIT: Monomer.
 CC -1- TISSUE SPECIFICITY: Expressed specifically in the hypopharyngeal
 glands of the forager (worker) honeybee.
 CC -1- SIMILARITY: Belongs to family 13 of glycosyl hydrolases.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation
 at the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: D79208; BAA11466.1; -.
 DR PIR: UC4714; UC4714.
 DR HSSP: P21332; 1U0K.
 DR InterPro: IPR006589; Alp_amy1_cat_sub.
 DR InterPro: IPR006047; Alpha_amy1_cat.
 DR Pfam: PF00128; alpha-amy1ase; 1.
 DR SMART: SM00642; Amy1; 1.
 KM Hydrolyase; Glycosidase; Glycoprotein; Signal.
 FT SIGNAL 1 17
 FT CHAIN 1 567
 FT CARBOHYD 88 567 ALPHA-GLUCOSIDASE.
 FT CARBOHYD 123 123 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 247 247 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 290 290 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 313 313 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 319 319 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 499 499 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 507 507 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 567 AA; 65565 MW; A148BD657G99C2 CR664;
 Query Match 9.6%; Score 325.5; DB 1; Length 567;
 Best Local Similarity 21.6%; Pred. No. 2-2e-17;
 Matches 128; Conservative 86; Mismatches 198; Indels 181; Gaps 23;
 QY 118 GDLKGLKRIKIPYFQGLTLYLIMPLFKCEKSGSGVAVSSRVNPAALGTGDLREVT 177
 DB 49 GDIEIKRKLDHFLKLEMGVDMFWLSPLY--PSFMVDFGDISINTYVHPIFGISLDLNV 106
 QY 178 AALHAGISAVVDITFNHTSNEHEWAQRCAAGDPLFDNFYITP-----DRMPDYDRT 232
 DB 107 SAAHEKSKJLITLDFPNHTSDQHEWFOJSLKIKIEYNNYYIWHPKIYNGKVP----- 160
 QY 233 LREIPDQHPGFGFSGLEGGRWWT-----TNSFGMDLNTYNSPNVFRMAEMFL 283
 DB 161 -----PTWVGVEFG--GSAMSWREERQAYTYIHQFAPEOPDLNTYNNPVLLDMQNVLRW 212

QY 284 ANLGVDILMDAVAFVWKOMTSCSE-----NLPOAHALIRAFNA 322
 DB 213 LRKGFDERVDALYICEDMRFLDEPLSGETNDNKTEYLIKITHDIPETVYVNRKRD 272
 QY 323 VWRIAAPVPEKSAIVHPDVOY--IGDECOIGNPLOMALLMNTLATREYNLHOA 380
 DB 273 VLDFEPQPKHMLIAYNTLSMTKRYDYGAD-----FPFNFAPIKN--VSRDSN---S 320
 QY 381 LTYEHNLPEHIAWNVYVRSHDDIGWTFPDEDAAYIGISGYHROKLPNFVYRFGSGFAR 440
 DB 321 SDFKLVYDN--WMTYPPSGIPWVPGNHQRLR-----VSRF----- 356
 QY 441 GVPEQVNSTGDCRVSGAALVGLAODDPHAYVRKILLYSLATSGGLPLIYLDVEVG- 499
 DB 357 -----GEEKARITTMSSLLPQVAVVYDEDEIGM 385
 QY 500 -----TLND-----DWSQDSN--KSDSRWAHPRRYNE 526
 DB 386 SDYIISWEDTQDPQCGAGKENYQTMRSRDPARTPFQMDSVSAGRSSSNTWL--RVNE 442
 QY 527 -----ALYAQRNDSTAGQIYQGLRHLIAVRSRPRDGGRLVTFNTN--NKHIIQYIR 579
 DB 443 NYKTVNLAAEKDKNS-----FFNMFKKFASLKSPYKEANL--NTRMLDNVFAFSR 494
 QY 580 -----NNALLAFGNFSEYPOVTAHTLQAMPFAHDLIGG-----KTVSLNQDL 623
 DB 495 ETEBNGSLYALINTSNEQIYDLKAFNNVPKKLNFFNNFNSDIKSINNEGV 547
 RESULT 7
 MGTA THEME
 ID MGTA THEME STANDARD; PRT; 441 AA.
 AC P80059;
 DT 01-AUG-1992 (Rel. 23, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE 4-alpha-glucanotransferase (EC 2.4.1.25) (Amylomaltase)
 DE (Disproportionating enzyme) (D-enzyme).
 GN MGTA OR TM0364.
 OS Thermotoga maritima.
 OC Bacteria; Thermotogae; Thermotogales; Thermotogaceae; Thermotoga.
 CX NCBI_TaxID=2336;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MSB8 / DSM 3109 / ATCC 43589;
 RX MEDLINE=99287316; PubMed=10360571;
 RA Nelson K.E., Clayton R.A., Gill S.R., Gwinn M.L., Dodson R.U.,
 RA Haft D.H., Hickley E.K., Peterson J.D., Nelson W.C., Ketchum K.A.,
 RA McDonald L., Utterback T.R., Malek J.A., Linher K.D., Garrett M.M.,
 RA Stewart A.M., Cooton M.D., Pratt M.S., Phillips C.A., Richardson D.,
 RA Heidelberg J., Sutton G.G., Fleischmann R.D., Eisen J.A., White O.,
 RA Salzberg S.L., Smith H.O., Venter J.C., Fraser C.M.;
 RT "Evidence for lateral gene transfer between Archaea and Bacteria from
 RT genome sequence of *Thermotoga maritima*.";
 RL Nature 399:323-329(1999).
 RN [2]
 RP SEQUENCE OF 1-23, AND CHARACTERIZATION.
 RC STRAIN=MSB8 / DSM 3109 / ATCC 43589;
 RX MEDLINE=92331687; PubMed=1628664;
 RA Liebl W., Fell R., Gabelsberger U., Kellermann J., Schleifer K.-H.;
 RT "Purification and characterization of a novel thermostable 4-alpha-
 RT glucanotransferase of *Thermotoga maritima* cloned in *Escherichia
 RT coli*.";
 RL Eur. J. Biochem. 207:81-88(1992).
 RN [3]
 RP X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS).
 RX MEDLINE=22135897; PubMed=12139940;
 RA Rouleauikova A., Raasch C., Sedelnikova S., Liebl W., Rice D.W.;
 RT "Crystal structure of *Thermotoga maritima* 4-alpha-glucanotransferase
 RT and its acarbose complex: implications for substrate specificity and
 RT catalysis.";
 RL J. Mol. Biol. 321:149-162(2002).


```

QY 223 LREIPDHPGSGSLED--GHWVTTNSFQMDLNTSPNVPFRAMAGMLFLANLGYDI 290
DB 142 --SVF--SGSAMVEDMTGEYTLHFSKQPDINWENPKYRREYEMKFWMLDKGVGD 195
QY 291 LAMDVAPIWK-----QMGTSGE-----NLPOHALIRAFNAVMRAADAVPFX 334
DB 196 FPMDDVIMNISKPELPDGPQSGKYYASGRYYMNGPRVHEFLQENK----- 242
QY 335 SEAIHVPDQVQYIGDECOIGYNPLQWALLMTLATREVNLLHQAITYRHNLPEHT--- 391
DB 243 REVLASKYD--IMTVGETP--GVTPKE--GILYTDSSRRRLNANVPQF-----EHMDLD 288
QY 392 -----AMVNYVRSHDDIGWTFAD-----EDAAV--LGISGVDRHOFNLRFVFN 432
DB 289 SGPQGWK-----DIRFWSLADLKKTMTKQKELEGKGMNSLYLNHNDQPAVSERF-- 338
QY 433 RPDGSGFAGVPPQNPSTGDCVSGTAALVGLADDDPHAVRILKLYSIALSTGSLPLI 492
DB 339 -----GD-----DGYRVSALQMLATFTLHMOSTPYI 365
QY 493 YLGDVGVG-----TLN-----DDWSQDSNKS----- 513
DB 366 YGGERIGMTNVPFSPEDRDLETILNMYKERYEVEGDPEQVMEKITYYGRNARTPMQW 425
QY 514 DDSR-----WAA-RPRYNEA-LYAQRNDPSTAAGIYQGLRHMTAVRQSNPREDDG 562
DB 426 DSENAAGTAGTPWIPVNPENYKEINVKALIEDPNS--VFHYKKLIDLRKQHIIVYG 481
QY 563 RLVYNTNNKATIGYIR--NNALLAFGNSFEPVQVNTANTLQAMPFKAHDLIGGTVSL 619
DB 462 TYDILLDEDDPYIRYTRITLGNELVITNFSF-----KTPVFLPDHIIYKT--- 528
QY 620 NQDLTLQPYQV 630
DB 529 --KELIISNDV 538

RESULT 9
MGTA THEME STANDARD; PRT; 442 AA.
ID MGTA THEME STANDARD; PRT; 442 AA.
AC 086956;
DB 30-MAY-2000 (Rel. 39, Last sequence update)
DB 30-MAY-2000 (Rel. 39, Last sequence update)
DB 10-OCT-2003 (Rel. 42, Last annotation update)
DB 4-alpha-glucanotransferase (EC 2.4.1.25) (Amylomaltase)
DB (Disproportionating enzyme) (D-enzyme).
DB MGTA.
DB Thermotoga neapolitana.
DB Bacteria; Thermotogae; Thermotogaceae; Thermotoga.
DB NCBI_TaxID=2337;
DB [1]
DB SEQUENCE FROM N.A.
DB STRAIN=22706-MC24;
DB Zverlov V.V., Maichenko O.V., Liebl W., Velikodvorskaya G.A.,
DB Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.
DB -1- CATALYTIC ACTIVITY: Transfers a segment of a (1,4)-alpha-D-glucan
DB to a new 4-position in an acceptor, which may be glucose or (1,4)-
DB alpha-D-glucan.
DB -1- COFACTOR: Binds 1 calcium ion per subunit (By similarity).
DB -1- SUBUNIT: Monomer (By similarity).
DB -1- SUBCELLULAR LOCATION: Cytoplasmic.
DB -1- SIMILARITY: Belongs to family 13 of glycosyl hydrolases.
DB This SWISS-PROT entry is copyright. It is produced through a collaboration
DB between the Swiss Institute of Bioinformatics and the EMBL outstation
DB the European Bioinformatics Institute. There are no restrictions on its
DB use by non-profit institutions as long as its content is in no way
DB modified and this statement is not removed. Usage by and for commercial
DB entities requires a license agreement (See http://www.isb-sib.ch/announce/
DB or send an email to license@isb-sib.ch).
DB EMBL; AJ009831; CA08864.1; -
DB InterPro; IPR006569; Alp_amy1_cat_sub.

```

```

DR InterPro; IPR006047; Alpha_amy1_cat.
DR InterPro; IPR006046; Glyco_hydro_13.
DR Pfam; PF00128; alpha-amy1ase; 1.
DR PRINTS; PR00110; ALPHAAMY1ASE.
DR SMART; SM00642; Amy; 1.
KW Carbohydrate metabolism; Transferase; Glycosyltransferase;
KW Calcium-binding.
FT ACT_SITE 186 186 BY SIMILARITY.
FT ACT_SITE 216 216 BY SIMILARITY.
FT ACT_SITE 278 278 BY SIMILARITY.
FT METAL 13 13 CALCIUM (BY SIMILARITY).
FT METAL 15 15 CALCIUM (BY SIMILARITY).
FT METAL 17 17 CALCIUM (BY SIMILARITY).
FT METAL 19 19 CALCIUM (VIA CARBOXYL OXYGEN) (BY
FT METAL 21 21 SIMILARITY).
SQ SEQUENCE 442 AA; 51935 MW; AD42748834568C7 CRC64;

Query Match 9.0%; Score 308; DB 1; Length 442;
Best Local Similarity 25.6%; Pred. No. 3.5e-16;
Matches 114; Conservative 62; Mismatches 151; Indels 118; Gaps 19;

QY 107 VGVGYVDLF-----AGDLKGLDKPIPYFQELGLTYLHMLPKCEPKSGDGAYVSS 159
DB 2 IGQIYVSRFDGNDGVDGDFKGLKGLAISLYKELGVDFVWLMFV-----SSIFHGVDVD 58
QY 160 YEDVPALGTGDLREVIATLAEAGISAVVDFEINHTSNEHEMAQCAAGDPLFDNFYTI 219
DB 59 FYSFAEYGEDEKDFEMIEAFHNGIKVLDLPIHTHTGLFTWFOKALGDGPHYRD-YYV 117
QY 220 FPDKRPMPQYDRTLAIEFPDQ--HRGFSQLEDGRWTTTNSQMDLNTSPNVPFRAM 276
DB 118 WAS-----EKTDIDERRERDNERIWP-----LEDGRFTRGLRGPISPDNTNPOVEEM 168
QY 277 AGEMLFLANLGYDIAPMDVAFAIWKMGTSCEMLPOAHALIRAF-NAVMRIAPAVFEX 335
DB 169 KKVYVHLEMGVDGRFPA--KHMR--DTLQNVRFMRKYPIISDIEGIFLAELWAS 221
QY 336 EAIHVPDQVQYI--GDECOIGYNPLQWALLMTLATREVNLLHQAITYRHNLPEHTA 392
DB 222 KVDDEHGRIFGYMLNFDTSHC-----IKEA-VMKENFKVLESIERLAVGDYLP--- 270
QY 393 WNYVRSHDDIGWTFADDAALVLSGYDRHQLRFPNFRPDGSGFARGVPPQVNSTGD 452
DB 271 -VNFTSNND-----WKR--LASFEG----- 287
QY 453 CRVSGTAALVGLADDDPHAVRILKLYSIALSTGSLPLIYDGVG-----T 500
DB 288 -----GLSR-----EKVLSLSILFTLPGLPIFYGDGLGKMGKIRKPTVEYV 330
QY 501 INDDWSQDSNKSDDSRWAHPRYN 525
DB 331 IDPFPMSENISLHGQTFWKW-PAVN 354

RESULT 10
AM33 DICTH STANDARD; PRT; 498 AA.
ID AM33 DICTH STANDARD; PRT; 498 AA.
AC P14859;
DB 01-APR-1990 (Rel. 14, Created)
DB 01-APR-1990 (Rel. 14, Last sequence update)
DB 10-OCT-2003 (Rel. 42, Last annotation update)
DB Alpha-amy1ase 3 (EC 3.2.1.1) (1,4-alpha-D-glucan glucanohydrolase).
DB AM3.
DB Dicyogloms thermophilum.
DB Bacteria; Dicyogloms; Dicyoglomales; Dicyoglomaceae; Dicyoglomae.
DB NCBI_TaxID=14;
DB SEQUENCE FROM N.A.
DB STRAIN=H-6-12;
DB MEDLINE=88329076; PubMed=2458257;
DB Horinouchi S., Fukusumi S., Ohshima T., Bepu T.;
DB "Cloning and expression in Escherichia coli of two additional amy1ase
RT

```



```

OY 600 TLQAMPFAHDLIGKSTVSLNODLTLPORYOW 631
DB 527 EVELL-IHNYDVENGPI-----ENTTLRPEYAM 553

RESULT 12
MALI_DROME
ID MALI_DROME STANDARD; PRT; 508 AA.
AC P07191.
DT 01-APR-1988 (Rel. 07, Created)
DT 01-APR-1988 (Rel. 07, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Possible maltase D precursor (EC 3.2.1.20) (Larval visceral protein
D).
GN Drosophila melanogaster (Fruit fly).
OS Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
CX NCBI_Taxid=7227;
(1)
RN SEQUENCE FROM N.A.
RX MEDLINE=83216126; PubMed=6854639;
RA Snyder M., Davidson N.;
RT "Two gene families clustered in a small region of the Drosophila
genome."
RL J. Mol. Biol. 166:101-118 (1983).
(2)
RN POSSIBLE FUNCTION.
RX MEDLINE=88289361; PubMed=3135536;
RA Henikoff S., Wallace J.C.;
RT "Detection of protein similarities using nucleotide sequence
databases."
RL Nucleic Acids Res. 16:6191-6204 (1988).
CC -1- CATALYTIC ACTIVITY: Hydrolysis of terminal, non-reducing 1,4-
linked D-glucose residues with release of D-glucose.
CC -1- DEVELOPMENTAL STAGE: One of the proteins expressed by the 44d
cuticle gene cluster. Expressed in first, second and early 3rd
instar larvae and in adults, but not in embryos or pupae.
CC -1- SIMILARITY: Belongs to family 13 of glycosyl hydrolases.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation
at the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@sib-sib.ch).
-----
DR EMBL: V00204; CAA23492.1; -
DR PIR: S08597; S08597.
DR HSSP: P21332; 1UOK.
DR FLYbase; FBgn0002569; lvpD.
DR InterPro; IPR006589; Alp_amy1_cat_sub.
DR InterPro; IPR006047; Alpha_amy1_cat.
DR Pfam; PF00128; alpha-amy1ase; 1.
DR SMART; SM00642; Amy; 1.
DR Hydroxylase; Glycosidase; Signal.
KM SIGNAL
FT 1 23
FT CHAIN 1 23
FT CARBOHYD 30 508
FT CARBOHYD 124 124
FT CARBOHYD 198 198
FT CARBOHYD 312 312
SQ SEQUENCE 508 AA; 57941 MW; C0841DEB5A4DCEA CRC64;

Query Match 8.6%; Score 294; DB 1; Length 508;
Best Local Similarity 22.2%; Pred. No. 5.2e-15;
Matches 119; Conservative 83; Mismatches 205; Indels 128; Gaps 19;
OY 118 GDLKGLKDKIPYQELGLTYLHMLPKCPBGKSGGVVSSRYDNPALGTLGLREVI 177
DB 50 GDLKGTSHLGYLKEGTATWLSPTSP--MSPDGYISNFRYDIDPFGTLDEFDLLI 107

```

```

OY 178 AALHEAGISAVVDFEINHTSNEHEMAORCAAGDPLFDFNYFYIPDRRMDQYDRLEIF 237
DB 108 VEAKSLGKTLILDVFNHSSDENWVFEKSVNREIGYDF-YWPDGKINBTGAR----- 161
OY 238 PDQHGFGFSQLEDGRWVWT-----TFNSFGMDINTSNPWVFRAMAGMLFLANIGV 288
DB 162 -DPSNVWVSFSGPMWTFNNEKROQYFLHQGVKQPDINTPVRREHLDVLKFWLDRGV 220
OY 289 DILMDVAVAFWKMGKSGSCENLPQAHALIRAFNAVMRLAAVAFPKSAIYHPQVQYI 348
DB 221 DGFRIIDVAPVPIYHERNDGS-----YDEPPIVSGW 249
OY 349 GQDECOQGYNPLQALMLNTLATRE---VNLHQALYRHNLPENTAVWVYVSHDDIG 404
DB 250 GSDNNAYDYH-----DHYTQDPATVDMTEWREFFLDNYRAQNGDSRVLLAE--A 299
OY 405 WTPADEDAAVYLGISGYDHRQFLNFVFNRPDQ-SFARGV--PFOYNPSTGDCRVGTAA 461
DB 300 YSSVETLSAVFNGNSTHGTQLPNNFQMLYLSGYSTAKDVGSIDYWMWT-WMKHQYANW 358
OY 462 LVGLAQDDPRAVDR---IKLYSIALSTGGLFLIYGDVGTIUND----- 504
DB 359 VVG-NHDTNRVADMGAKHVDLLNVVALPAGAVTYGEBIGMSNVDECTGDSCEDRD 417
OY 505 -----DMSQDSN---XSQDSRMAH-RPRVNEALYAQNDSTAGQIYGLRHMIAVRQ 554
DB 418 GERTPMQWTKAKNADPSDGEETWLPSPRYRYVQVTRGVSRSLSINFKGLQELKS--- 474
OY 555 SNPRFDGRLVTFTNNKHIIQYIRNNALIAF---GNSEXPQYVTAHTLQAMPF 606
DB 475 -----SSAFIAFKEDGFSY--EAVTEQVLQIRY 502

RESULT 13
AMY_BACME
ID AMY_BACME STANDARD; PRT; 520 AA.
AC P20845;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Alpha-amy1ase precursor (EC 3.2.1.1) (1,4-alpha-D-glucan
glucanohydrolase).
DE Bacillus megaterium.
OS Bacillus; Firmicutes; Bacillales; Bacillaceae; Bacillus.
CX NCBI_Taxid=1404;
(1)
RN SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RX MEDLINE=88262518; PubMed=2455281;
RA Metz R.J., Allen L.N., Cao T.M., Zeman N.W.;
RT "Nucleotide sequence of an amy1ase gene from Bacillus megaterium."
RL Nucleic Acids Res. 16:5203-5203 (1988).
CC -1- CATALYTIC ACTIVITY: Endohydrolysis of 1,4-alpha-glucosidic
linkages in oligosaccharides and polysaccharides
-1- COFACTOR: Binds 1 calcium ion per subunit (By similarity).
CC -1- SUBUNIT: Monomer (By similarity).
CC -1- SIMILARITY: Belongs to family 13 of glycosyl hydrolases.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation
at the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@sib-sib.ch).
-----
DR EMBL: X07261; CAA30247.1; -
DR PIR: S01031; S01031.
DR HSSP: P21332; 1UOK.
DR InterPro; IPR006589; Alp_amy1_cat_sub.
DR InterPro; IPR006047; Alpha_amy1_cat.
DR InterPro; IPR006046; Glyco_hydro_13.
DR Pfam; PF00128; alpha-amy1ase; 1.

```

DR PRINTS; PRO0110; ALPHAAMYLASE.
 DR SMART; SM00642; Amy; 1.
 KM Carbohydrate metabolism; Hydrolase; Glycosidase; Calcium-binding;
 FT SIGNAL. 1 27
 FT CHAIN 28 520 ALPHA-AMYLASE.
 FT ACT SITE 233 233 BY SIMILARITY.
 FT ACT SITE 273 273 BY SIMILARITY.
 FT ACT SITE 340 340 BY SIMILARITY.
 FT METAL 140 140 CALCIUM (BY SIMILARITY).
 FT METAL 203 203 CALCIUM (BY SIMILARITY).
 FT METAL 237 237 CALCIUM (VIA CARBONYL OXYGEN) (BY
 SIMILARITY).
 SQ SEQUENCE 520 AA, 58761 MW, 206490575835C0 CRC64;
 Query Match 8.6%; Score 292; DB 1; Length 520;
 Best Local Similarity 22.8%; Pred. No. 7.6e-15;
 Matches 120; Conservative 74; Mismatches 208; Indels 124; Gaps 18;
 QY 112 YVDLF-----AGDLKGLKDKIPYF-----QELGLTYLHMLPLFKCPGKSDGGA 156
 DB 46 YVNSFYDANKGHGDLKGLTKDKLYNDGNSHTKNDLYQNGIWMFVNSP---SYHKID 102
 QY 157 VSSYSDVNPALGTIGDLREVIALHRAGISAVVDFIHNSTSEHEWAQRCAAGDPLFDFNF 216
 DB 103 VTDYINIDPQXNLDQFRKMKRQDVVIMDLVNTSTSEHEPWFQALDKNSKYRD 162
 QY 217 YVIFPDREMPDQYDRTLRIFPDQHPGFSQ-----LEGRVWTTTNSFQMDLANSNW 271
 DB 163 YYIMADK-----NTDLNE-----KSGWGOQVWHRAPGEYFYGTFWGMDPLNDPBE 210
 QY 272 VFRRAAGMELFLANIGVDILRMDAVAFIKQMGTSCEPLPOHALLIRAFNAVRIAPAV 331
 DB 211 YKEMINIVKFMKQGVDPFRDLALHTRK--GQTEGAKKNTLMWNEFRDAMKKNPV 268
 QY 332 FKESKALVHPDOVQYIGQDECOIGYNPLQMLMNTLATREVNLIHQALTYRHNLPBET 391
 DB 269 YLTGEWDPQEVAVP-----YQSLDSLFNFDLAG 298
 QY 392 AMVNTYRSDDIGWTTADEDAATLGISGDHROFLARFPVNRPDGSPAGAVPQVNPSPG 451
 DB 299 KIVSSVKAANDQ-----IATMAATBELFKSYNPNKIDGIFLT---NHQNRV 344
 QY 452 DCRVSTAAALVGLAODDPAVDRILKLYSIALSTGGLPLIYAGDVGTLLND--DD----- 505
 DB 345 MSELSD-----YKAKSAAISILLTLPFNPIYIGEIGMTGEPDLILBE 390
 QY 506 ---WSQDSNKSDDSRWHRPRYN-----EALVYQRNDPSTAGQIYQGLRHMLAVRQ 554
 DB 391 PPRW-YEGNGIGQTSW-ETPVYVKKGNQGVSEAOQTQKD-----SLNHYREMLRVQ 441
 QY 555 SNRPEDGRLVTENTNNKHIIIGTIR--NNALLAFGNFSEYPTVT 597
 DB 442 QHEELVKGTLQGISVDSKEVAVSRYYKSGISVYHNISNPVKVS 487
 RESULT 14
 ID 016G_BACCO STANDARD; PRT; 555 AA.
 AC 045101;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE 019G-1,6-glucosidase (EC 3.2.1.10) (Oligosaccharide alpha-1,6-
 glucosidase) (Sucrease-isomaltase) (Isomaltase) (Dextrin 6-alpha-D-
 glucanohydrolase).
 DE GN
 DE MALT.
 OS Bacillus coagulans.
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 OX NCBI_Taxid:1396;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 7050;

RX MEDLINE=96236197; PubMed=8787404;
 RA Watanabe K., Kitamura K., Suzuki Y.;
 RT "Analysis of the critical sites for protein thermostabilization by
 RT proline substitution in oligo-1,6-glucosidase from Bacillus coagulans
 RT ATCC 7050 and the evolutionary consideration of proline residues";
 RL Appl. Environ. Microbiol. 62:2066-2073 (1996).
 CC -1- CATALYTIC ACTIVITY: Hydrolysis of 1,6-alpha-D-glucosidic linkages
 CC in some oligosaccharides produced from starch and glycogen by
 CC alpha-amylase, and in isomaltose.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -1- SIMILARITY: Belongs to family 13 of glycosyl hydrolases.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.1sb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC EMBL; D78342; BAA11354.1; -.
 CC HSSP; P21332; IUOK.
 DR InterPro; IPR006589; Alp_amy1_cat_sub.
 DR InterPro; IPR006047; Alpha_amy1_cat.
 DR Pfam; PF00128; alpha-amy1ase; 1.
 DR SMART; SM00642; Amy; 1.
 KM Hydrolase; Glycosidase.
 FT ACT SITE 199 199 BY SIMILARITY.
 FT ACT SITE 255 255 BY SIMILARITY.
 FT ACT SITE 332 332 BY SIMILARITY.
 SQ SEQUENCE 555 AA, 64990 MW, CE2BBID7BAE226D8 CRC64;
 Query Match 8.5%; Score 289.5; DB 1; Length 555;
 Best Local Similarity 21.7%; Pred. No. 1.3e-14;
 Matches 133; Conservative 79; Mismatches 220; Indels 181; Gaps 23;
 QY 118 GDLKGLKDKIPYFOELGLTYLHMLPLFKCPGKSDGGA VSSYSDVNPALGTIGDLREV 177
 DB 27 GDLRGIMDLKTLKTLGIDICWISPYVDSQ--DNGYDIRDYRKIDKXGTEDMDRL 84
 QY 178 ALHRAGISAVVDFIHNSTSEHEW--AQRCAAGDPLFNFXYIFPDREMPD----- 227
 DB 85 DEAHARGIKTWMDLVNNTHSDHAWFVESRKKDNKYRD--FTWQDPK--PDGTPPNWG 141
 QY 228 -----QYDRTLRIFPDQHPGFSQLEDRVWTTTNSFQMDLANSNPWFVRMAGE 279
 DB 142 SNFSGAMEYDET-----TQYIHYTSKQPDLMWNEKRYKXIYK 184
 QY 280 MLFLANTGVDLIRMDAVAFIKR-----QMGTSCEPLPOHALLIRAFNAV 324
 DB 185 MKFWMQKGVDMRMVIGISKFLDPDYELPEGQKYGIGKYHANGPRLHAFIOEMNR-- 242
 QY 325 RIAARAVFKSALVHPDOVQYIGQDECOIGNPLQMLMNTLATREVNLIHQALTYR 384
 DB 243 -----EVLSKYDQMT-----VEIKRKTGPDREHLMNT----- 278
 QY 385 HNLPEHTAWNYVRSHDDIG-----WTEADEDAATLGISGDHROFLNFFVVRPD 435
 DB 279 -----FNF--KEMVDITRPGSPAGKALXPLVL-----KQLISMQVELAD 320
 QY 436 GSPFARGVPQPNPSTGDCRVSTAAALVELAODDPAVDRILKLYSIALSTGGLPLIY 495
 DB 321 TGM-NALYFENHDQ-----ARVSRWGDVITYRACAKAFATILGLKGTPIYQG 370
 QY 496 DEVGTLND-----DWS 507
 DB 371 EELGMVNALELEEDYDILIRNAYOELVVENQIMSODEPLTAVRKKGRDNARIPMWDGS 430
 QY 508 QDSNKSDDSRWHR--RPYNEALYACRNDPSTAGQIYQGLRHMLAVRQSNRPDGGRLVT 566
 DB 431 FNAAGFTTGTPMLKNSRYSIENVAKALQEPDSIFYYQSL---IKLRHSGVDTDRXYEL 487
 QY 567 FNTNNKHIIIGYIRNN--ALLAFGNFSEYPTVTANTHTQAMPKADLIGGKTVSLNQL 623

DB 488 IMPDHPHLYVYTRRESEKELVANAUSE--NTVSPD---QPDNKKLLGNEDTGTST 541
 QY 624 TLQPYO--VMMLE 634
 DB 542 LFRPYEALYYIE 554

RESULT 15
 DEXB STREQ
 ID DEXB STREQ STANDARD; PRT: 537 AA.
 AC Q59905:
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Glucan 1,6-alpha-glucosidase (EC 3.2.1.70) (Dextran glucosidase)
 DE (Exo-1,6-alpha-glucosidase) (Glucodextranase).
 GN DEXB.
 OS Streptococcus equisimilis.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxID=119602;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=H46A;
 RX MEDLINE=94049672; PubMed=8232196;
 RA Mechold U., Steiner K., Vettermann S., Malke H.;
 RT "Genetic organization of the streptokinase region of the
 RT Streptococcus equisimilis H46A chromosome.";
 RL Mol. Gen. Genet. 241:129-140(1993).
 CC -1- FUNCTION: The physiological substrates may be short
 CC -1- isomaltoosaccharides (By similarity).
 CC -1- CATALYTIC ACTIVITY: Hydrolysis of (1->6)-alpha-D-glucosidic
 CC linkages in 1->6-alpha-D-glucans and derived oligosaccharides.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -1- SIMILARITY: Belongs to family 13 of glycosyl hydrolases.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: X72832; CAA51348.1; -
 DR PIR: S39970; S39970.
 DR HSSP: P21332; 1UOK.
 DR InterPro: IPR006589; Alp_amy1_cat_sub.
 DR InterPro: IPR006047; Alpha_amy1_cat.
 DR Pfam: PF00128; alpha-amy1ase; 1.
 DR SMART: SM00642; Amyy; 1.
 KW Hydrolyase; Glycosidase.
 SQ SEQUENCE 537 AA; 61733 MW; 154DF0ACAF302FC7 CRC64;

Query Match 8.5%; Score 289; DB 1; Length 537;
 Best Local Similarity 22.0%; Pred. No. 1.4e-14;
 Matches 127; Conservative 89; Mismatches 236; Indels 124; Gaps 23;

QY 118 GDLKGLKDIPIYFOELGLTYLHMLPFCPEKSGGYAVSVYDVNPAIGTIGDLREVI 177
 DB 28 GDLKGLTTSQDDYIKGLGTRAIWLSFYQSP--MDNGYDISDYEAIAEYFGAMDDMDLL 85
 QY 178 AALHEAGISAVVDIFINHTSNHEMAQRCAAGDPLFDNFYIIPDRMPDQYDRITREIF 237
 DB 86 AANENGIKIMDLVYNHTSDEHAFVEARENPNSPERDYIWRDE-----PNNIMSIIF 139
 QY 238 PDQHPGSGESQLEB--GRWWTTFNSFQNLINSNPWFPRAMAGEMFLANLGVDILRMKA 295
 DB 140 ---SGSAWELEDAASQYYLHLFSKQPDLMNENAFVROKIYDMNMFIAKGIIGGRMDV 195
 QY 296 VAFIYWKQMGVSCSE--NLPOAHALIRAFNAVMRIAPAVFPEKSEALVHPDQVVOYIGQDECO 354

DB 196 IDLIGKIPDSEITNGPRRIHDYLKEMN-----QATFGNHD-----VMTVGE--T 237
 QY 355 IGYNPLOMALIMNTIATREVNLL---HQAITYRHNLDEHTAMVYVYRSHDDIGTFADE 410
 DB 238 WGATP-EIARQYSRPENKEISMVQFEBHVGLQHKPNAPK-----WDYAE 281
 QY 411 ---DAVIGISGYHROFL---NRFVNRD-----GS-----F 438
 DB 282 LDVPAKTFESKWOTELKLGEMNSLENNHDLPRVLSIWGNDSTYREKSAKALAILHL 341
 QY 439 ARGVPQYNPSTGDCRVSGTAALVGLAQDDPHAVDRIKLL-YSLALSTGGLPLIYIGDE 497
 DB 342 MRGTFPIYI-----QGEIIGMTNYPFKQLTSEVDLIESLNTYKAMENGVPARVWSS 392
 QY 498 VGLTNDLDD-----WSQDS---NKSDDSRNAHRPRYNEALYQQRNDPSTAGQIYQGLR 547
 DB 393 IRKVGDRNARTPMQWSKTHAGFSEAGETWLPVNPVYQEIINVA---DALANDOSIFYTYQ 449
 QY 548 HMAVROGNSPRFDDGRVLTFTNKKHLLGYIRNNALLAFGNPSEYQCTVTAHTLQAMPEK 607
 DB 450 QLIARKKQDWLVEADYHLLPFADK-VYAYOR-----QGEETIYIVVNVSDQEOVF 500
 QY 608 AHDLIGKTVSLNODLT-----LQPYQVMMLEIA 636
 DB 501 AKDLAAGAEVLTNTDVEDKVLKTHLQPMWDAFCVKLS 536

Search completed: May 25, 2004, 19:31:18
 Job time : 43 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Inc.

OM protein - protein search, using SW model

Run on: May 25, 2004, 19:24:11 ; Search time 20 Seconds
(without alignments)
3058.891 Million cell updates/sec

Title: US-09-807-146-1

Perfect score: 3408
Sequence: 1 MLTPQGVGLIQLYIKTRIL.....VSLNQDLTLQPYQVWMEIA 636Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 200000000Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summariesDatabase : PIR 78:*
1: PIR1:*
2: PIR2:*
3: PIR3:*
4: PIR4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1319.5	38.7	644	2 C75457	alpha-amylase - De
2	919.5	27.0	584	2 C87390	amylase, alpha-1,4
3	541.5	15.9	1100	2 G83376	probable trehalase
4	520.5	15.3	552	2 E75322	probable trehalase
5	509.5	15.0	601	2 G70983	probable glycogen
6	504	14.8	573	2 S71450	trehalase synthase
7	452.5	13.3	568	2 H64879	probable membrane
8	445	13.1	544	2 C95854	Sucrose phosphoryl
9	393	11.5	566	2 AF1803	probable phosphoryl
10	386.5	11.3	434	2 B85754	Sucrose phosphoryl
11	373	10.9	560	2 AF1416	Sucrose phosphoryl
12	333.5	9.8	548	2 AF1231	alpha, alpha-phosph
13	330.5	9.7	549	2 S44188	alpha-glucosidase
14	325.5	9.6	441	2 S60618	4-alpha-glucanotri
15	325.5	9.6	567	2 JC4714	alpha-glucosidase
16	321.5	9.4	549	2 C89930	alpha-D-1,4-glucos
17	318.5	9.3	548	2 AB1585	alpha, alpha-phosph
18	314	9.2	541	2 A41707	oligo-1,6-glucosid
19	309.5	9.1	562	2 H96083	dextranase (EC 3.2
20	307	9.0	587	2 S42358	alpha-glucosidase
21	304	8.9	498	2 S01313	alpha-amylase (EC
22	304	8.9	558	2 S13579	oligo-1,6-glucosid
23	302.5	8.9	564	2 H75403	glycosyl hydrolase
24	297.5	8.7	541	2 B95220	dextran glucosidas
25	295	8.7	546	2 F70866	hypothetical maltase
26	294	8.6	508	2 S08597	hypothetical prote
27	293.5	8.6	520	2 S55363	maltase-like prote
28	292	8.6	520	2 S01031	alpha-amylase (EC
29	289	8.5	537	2 S39970	glucan 1,6-alpha-g

30	288.5	8.5	562	2 C92263	trehalase-6-phosph
31	287.5	8.4	498	2 S55362	maltase-like prote
32	286.5	8.4	561	2 AE2649	alpha-glucosidase
33	286.5	8.4	594	2 D97431	probable alpha-glu
34	286	8.4	555	2 AH0449	alpha, alpha-phosph
35	285	8.4	522	2 S07253	hypothetical prote
36	285	8.4	534	2 T30268	oligo-1,6-glucosid
37	281	8.2	505	2 S08598	hypothetical prote
38	281	8.2	584	2 S46183	alpha-glucosidase
39	280.5	8.2	551	2 H91280	trehalase 6-P hyd
40	280.5	8.2	551	2 H86121	trehalase 6-P hyd
41	280	8.2	584	2 ALBY	alpha-glucosidase
42	280	8.2	584	2 S64627	alpha-glucosidase
43	279.5	8.2	550	2 AB1058	alpha, alpha-phosph
44	278	8.2	553	2 A11097	oligo-1,6-glucosid
45	277	8.1	553	2 AH1460	oligo-1,6-glucosid

ALIGNMENTS

RESULT 1

C75457
alpha-amylase - Deinococcus radiodurans (strain R1)
C:Species: Deinococcus radiodurans
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 17-Mar-2000
C:Accession: C75457
R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, B.K.; Peterson, J.D.; Dodson, R.J.;
M.; Shen, M.; Vamathevan, J.D.; Lam, P.; McDonald, L.; Uterback, T.; Zalewski, C.; Mc
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
Science 286, 1571-1577, 1999
A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
A:Reference number: A75250; MUID:20036896; PMID:10567266
A:Accession: C75457
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-644 <NHI>
A:Cross-reference: GB:AE001946; GB:AE00513; NID:96458655; PIDN:AA10510.1; PID:9645865;
A:Experimental source: strain R1
C:Genetics:
A:Gene: DR0933
A:Map position: 1
C:Superfamily: alpha-glucosidase; alpha-amylase core homology

Query Match	38.7%; Score 1319.5; DB 2; Length 644;
Best Local Similarity	42.1%; Pred. No. 8e-96;
Matches	281; Conservative 110; Mismatches 215; Indels 61; Gaps 14;
QY	1 MLTPQGVGLIQLYIKTRILDIYTPQKAGTEKSEDMQFSRHMTHPKMLNGLSDVYG 60
DB	1 MLTPD-----LAARL-----RLAFDDRDARTFLRLRYGPELDADNIRAYIG 43
QY	61 NN-EALTL-PMLEMLLAQMSYQSNSSLKIDIRENNPPWILSNKQGVGYVDLFAIG 118
DB	44 NHAADLIGELIRKVL-----HAYHARPADIKLDEARLRPLQLQPEWVGYAYVDRFAG 99
QY	119 DKGKDKIPYFOELGLTYLHMLPKCPCKEKSQGVAVSSYRDVNPALGTTGIDREYIA 178
DB	100 TLKRGGERLVEYIEGLGVYTLHLPLLRPDENDGAVQYRSVRPLGTTIDDLASLAR 159
QY	179 ALHENGISAVVDFTFNHNSNEHMQRCAGDPLEFDNYVYFFPDRMDQVORTLRETFP 238
DB	160 ELRGGRISLVLDLVNHAEEHBMVVRATAGAAARDYFHLFPDTPQPAERILPELFP 219
QY	239 DQHPGFGS-QLEDGRWVTFNFSFQMDLNVSNPWFARAGMELFLANIGVDILMDAVA 297
DB	220 DPAPGNFTWNGRAGGWVTTFNRSQMDVNWGNPAVFRFYLDLITLARGVVFRDLAIA 279
QY	298 FTKWMGTSCEMLPOAHLIDAFNAVMRIAPAVFFKSDATVHPQVQVYIGQDE----- 352
DB	280 FLMKRLGTDCQNGQVEVRLTHALRAATRIVAPAVAFKAIAVAPDGLIHYIGSSRDHGRV 339
QY	353 CQIGYNPLQAMLLMNTLATREVNLIHQALYTRHNLPHRTAWNVYVRSHDDIGMTFADEDA 412

Db	340	SDMAAHNSLVQWMSLASRDRLRLPALAAAPPKPTNTMGVYVRCHDDIGMAINDEDA	3599
QY	413	AYIGISGYDHRQVLANFFVNRPDGSPARAVPQINPSTGDCRVSTAAVLGL-----	4655
Db	400	ARVGLSGPARHRLSPFYSGEPFGSPARGLVQHHPQTDGDRRISGTAASLAGDLALLETG	4599
QY	466	-AODDPHAYDRILVLSIALSTGGRLYLIGSEVSTLDDDDWSQDSNKSDDSPRAHPRY	5244
Db	460	DAERVDALLARLLLLHAAVMLGFGGVPLTMSGELLALNDTPAAVPAHAPDAVRVHRPQM	5199
QY	525	NEALYAQ-----RNDPSTAAGIYOGLRHMLTAVRQSNRPREDGRLVTFNTNKH-----	5744
Db	520	DWELVASAQADATGQPVTPAGMFMFGRLHLLAAVRRTRPHLHA-----STESKPLRSPD	5733
QY	575	--IGYLR---NNALLAFGNSESYEPQVTAHTLQ--AMPKKAHDLIGKTVSL--NODILLQ	6266
Db	574	PCVILLRRREHPGVLLQVNVFSEHHHTTPTPMLQEQGLAVAHDLGSEQFTHGSPDLALE	6333
QY	627	RYQVMML 633	
Db	634	PYRALWL 640	

RESULT 2

```

C87390      amylosucrase [imported] - Caulobacter crescentus
C|Species: Caulobacter crescentus
C|Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 20-Apr-2001
C|Accession: C87390
R|Niernan, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.
B.; Laub, M.T.; Deboy, R.T.; Dodson, R.U.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon
n, V.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, U.C.; Fraser, C.M
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A|Title: Complete Genome Sequence of Caulobacter crescentus.
A|Reference number: A87249; MUID:21173698; PMID:11259647
A|Accession: C87390
A|Status: preliminary
A|Molecule type: DNA
A|Residues: 1-584 <STO>
A|Cross-references: GB:AE005673, NID:G13422449, PIDD:AAK23119.1, GSPDB:GN00148
|Genetics:
|Gene: CC1135

```

Query Match	27.0%;	Score 919.5;	DB 2;	Length 584;
Best Local Similarity	35.2%;	Pred. No. 1e-65;		
Matches 215; Conservative	100;	Mismatches 246;	Indels 49;	Gaps 16

QY 44 MDTHFFPLKMLNLDSSVYGNNEALLPMLLEMLLAQMSYSGRNSLTKDIDIAEENNPWILS 103
 Db 1 MESRFAPKL-----YGDARGPVAVIERKTMLLKQARAREPEPLRALDAAAPALMLHA 53
 QY 104 NKQVGVCCYVDLFAEDLKELKDKIPYEQELGLTYLHMLPLEKCEPKSGDGYAVSYRDV 163
 Db 54 PGQATATFYVDREFRAGDLNVRGKLDYLTDELGYRMHLPLLEPRPGSDGSGFAVADRVKY 113
 QY 164 NPLAGLTGDLREVIATLHEAGISAVVDFENHTSNEHMAORCAAGDLPENFYFIPDDR 223
 Db 114 DPLKLTIDDLDEALAGDLRQPDMLIDVVCNHTAEHMAAKARAGDPAVDYDTIYV.PDA 173
 QY 224 RMDQDYDRTLEIFPDQHPGGS-QLEDGRWWTTFNSFQDNLANSNPFVERAMAGMLTF 282
 Db 174 QSAARDRELIDVEPDTLPGSFYYDAAMGCVMTTFYFPQDMLNANPAVFAEMLEVLIF 233
 QY 283 LAILGVDLILMDVAITWKQMGTSCENLFOAALLRANVAWRLLAARVFFKSRLYHPD 342
 Db 234 LAAGKAGQFPLDAPFLKMKQAGTTCNPLFQYREIVEARPAALISTVAPGVLLAAELISVZ 293
 QY 343 QVQVYIGODE--QIIGNPFOALLNNTLATREVNLLHQALTYSRHLPEHTAWNVYRSH 400
 Db 294 DVLPFRGEGSSGGLNANNVMTALMAALADDDAVIARCLAVARKRAQAGAMLNYYRCH 353
 QY 401 DDIQGTFADEDAAYILGISGYDHRQPLNRFVNRFDGSPARGVPOYN---PSTGDCRVS 456

```

Db      354  DDLIMNAL--AAVAPAS--DLRBSNMY--GNGEG--FSRGRAFQTAEGGVSEF-----N
Qy      457  GTRAAALVGLAOD--DPHAYNRILLYALSTGLPLITYGDEVGTLINDDDMQSDNSK 51.3
Db      401  GNAALAGLTADAGCGTCGARRRLLYGIIHALDQWPLITMGDEIGLDNDEAYDDPLRA 46.0
Qy      514  DDSRMARHPRYNEALYAQNDPSTAGQIY-----QGLRHMTAVRQSNRPDGGFLYT 56.6
Db      461  GGRQWTHRRPQOMWSLAERRGEAGLQADLPATFARLGGARRRLTLTGAVG---APPV- 51.5
Qy      567  FNNNNHIIIGYIRNNA---LTAFGNSEYPTVAHTLQAMPFAHDLIGGKTVSLNQDL 62.3
Db      516  EVSSPAVLAFLARDGARPFLCVANVSDAPQDF--ELPPAPAQGAEVDLDG--AESPAGAV 57.1
Qy      624  TLQPIQVMTL 63.3
Db      572  SLPPYGITWL 58.1

```

RESULT

G83376
 probable trehalose synthase PA2152 [imported] - Pseudomonas aeruginosa (strain PA01)
 C/Species: Pseudomonas aeruginosa
 C/Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
 C/Accession: G83376
 R/Stover, C.K.; Pham, X.Q.; Errvin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; B
 adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kae, A.; Lardig, K.; Lam
 .; Lory, S.; Olson, M.V.
 Nature 406, 959-964, 2000
 A/Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic path
 A/Reference number: AB2350; MUID:20437337; PMID:10984043
 A/Accession: G83376
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-1100 <STO>
 A/Cross-references: GB:AE004642; GB:AE004091; NID:59948163; PIDN:AG05540.1; GSPDB:GN00
 A/Experimental source: strain PA01
 C/Genetics:
 A/Gene: PA2152

Query Match	15.94;	Score 541.5;	DB 2;	Length 1100;
Best Local Similarity	26.94;	Pred. No. 6.3e-35;		
Matches 165;	Conservative 90;	Mismatches 22;	Indels 135;	Gaps 25

[illegible]

QY 485 STGGLPLVLGDDEVT-----LNDDD-----WSQSDN-----KSDSGRMHRRPNEAL 528
 Db 379 SMPGPIPLVYDEIGMGDNITLIGDRDGVRTPMQMSVDRNGGFSRADDPKLV-LPPIIDPL 437
 QY 529 Y-----AQRNPSTPAQGIYQGLRHMIAROSNPRFDGRLVTENTNNKHIGYI-- 578
 Db 438 YGYQTINVEAQRDPHS---LNMNRRLLAVRSQKAFGRGSLMLLAPSRRIIAYLRE 493
 QY 579 -----KNNALLAFGNFSEYFQTVTAHTLQAMPFAH-----DLIGKTV---SLNDL 623
 Db 494 YAEGERQDSIIICVANTSLPAAQAV-----EIDLASHAGKVEVEMIGWSPPIGELTYLL 547
 QY 624 TLQRYQVWMLLEIA 636
 Db 548 TLPPYGFYWEYLA 560

```

Db      413 GGFSTAGPSCDFPPPIQDPVYIGRGVRNVQSOLDPPSS-----LTKTARQLRLRAHPAFA 468
QY      561 GGRLVTFENTNNKHIIIGYIR--NNALLAFGNFSBHYPTVTAHTLQAMPFKAHDLIGKTV 617
Db      469 HDLDTFTIEITGNPAIIAFTROYDETLIIYNSFAGNAO---AGLIDLAPF-----VGRAPV 520
QY      618 SLN-----QDLTIQPYQVMWLEI 635
Db      521 TLISGASPLPVTGNGQYPVVMGKIDYIMRL 551

RESULT 5
G70983
probable glycosyl hydrolase - Mycobacterium tuberculosis (strain H37Rv)

```

RESULT 4

E75322

Probable trehalose synthase - Deinococcus radiodurans (strain R1)

C:Species: Deinococcus radiodurans

C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 17-Mar-2000

C:Accession: E75322

R:White, O.; Eissen, J.A.; Heidelberg, J.F.; Hickey, E.R.; Peterson, J.D.; Dodson, R.J.; M.; Shen, M.; Vamathevan, J.T.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; MacScience 286, 1571-1577, 1999

S.: Smith, H.O.; Venter, J.C.; Fraser, C.M.

A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.

A:Reference number: A75250; MUID:20036896; PMID:10567266

A:Accession: E75322

A:Molecule type: DNA

A:Residues: 1-552 <WHI>

A:Cross-references: GB:AB002040; GB:AB000513; NID:g6459824; PIDN:AA11586.1; PID:g6459824

A:Experimental source: Strain R1

C:Genetics:

A:Map position: 1

C:Superfamily: alpha-glucosidase; alpha-amylase core homology

Query Match 15.3%; Score 520.5; DB 2; Length 552;

Best Local Similarity 26.4%; Pred. No. 1.1e-33;

Matches 151; Conservative 91; Mismatches 226; Indels 103; Gaps 18;

Qy 118 GDLGDKDKIYPFQELGTLTLMLPFEKPEGSGKDGAVSSYRDVNPALGTIGDLREVI 177

Db 31 GDFPGTISRRLDYLNKLGVDCLMLPWF--PGLADDDGDVADYGHFDLGTDDPKVFL 88

Qy 178 AALHEAGISAVVDITFNHTSNEHEWAQRCAGDPLD-----NFIYTFPPRRRPDQYDR 231

Db 89 REAHARGIRLVGDIVYNTHTSSDHPFQARRGPTLPQSGPNEYDYIYWSDE--GKGYAD 146

Qy 232 TLRETFPQHGEGFSQLED--GRWVTTFNFSFQOMDLANSMPVFRPAMGEMFLANTLGVDI 290

Db 147 T-RITFDTEVSNNITLDBQAKYYWHRFPAQDPDLNDNPKVYVELHGAARFMDLIGDG 205

Qy 291 LRMDAVAFWKQMGTSCENLFOAHLIRAFPAVWRIRAPAVFFSEALVHDDVQYQYIG- 349

Db 206 FRVAVAPTLIREGTSCENLEPETHILKGFRAWDREYFGRLTLEANNQWEEVEVEYGT 265

Qy 350 --QDECOGYNPLQMALINLTATREVNLLQALTYLTHNLEPHTAWVYVYSHDDICWTF 407

Db 266 EAEDEFHMCFFPMPFRIMYSLKREDDTSITREINGRLPKIIPSPQWCFFLNHDELTEM 325

Qy 408 ADEDAAYIGISGYDHROFLNRFVNNRPDGSFARGVPPOYNSTGDCRVY--GTAALVVG 464

Db 326 VTDDR-----AFMYAAYADARAKNVGIRRLAP 356

Qy 465 LAODPEHAVDRIKLITYSTALSTGGPLITYLDEVDGLTNDWSDSN--KSDDSRW----- 518

Db 357 LLDNDR---RIELNLTNYLALPGSPITYYDEIG--WDDDLGLPDRGVRPRPMQNNAGTS 412

Qy 519 -----AHRPRYNEALY-----AQRNDPSTAAQIYQGLRHMIAVROSNRPFD 560

C:Accession: G70983
 R:Colle, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S. Nature 393, 537-544, 1998
 A:Authors: Squares, R.; Sulistion, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G. A:Title: Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome A:Reference number: A70500; MUID:98295987; PMID:9634230
 A:Accession: G70983
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-601 <COL>
 A:Cross-References: GB:Z96071; GB:AL123456; NID:g3242254; PIDN:CAB09454.1; PID:G2181968
 A:Experimental source: strain H37Rv
 C:Genetics:
 A:Gene: Rv0126
 C:Superfamily: alpha-glucosidase; alpha-amylase core homology

Query Match	15.0%;	Score 509.5;	DB 2;	Length 601;
Best Local Similarity	27.6%;	Pred. No. 9,7e-33;		
Matches 160;	Conservative 83;	Mismatches 216;	Indels 121;	Gaps 24;
QY	117	AGDLGLKDKIKYFQELGLTYLHLMPLEKCEGKSGDYAVASRYDVNPAALGTIGDLREV	176	
DB	65	SDDLGLDLRLYLQWLGLDICIWLPPFYDSP--LRDGGIDLRDQYKYLPEEGTVDDFVAL	122	
QY	177	IAALHEAGISAVVDPIFNHTSNEHWAQ--RCAAGDPLFDFNYIIFPD--RRMPDQYRT	232	
DB	123	YDAARRRGIRITFDLVNHTSESHPFQESRRDPDGEYGD--YYWSDTSERYTD-----	175	
QY	223	LRKIFPDQHGFGSOLEDRW-----VWTTTNSFQMDLANSPPVFFAMAGEMLEFL	283	
DB	176	AKIIFVD-----TEESNWSFDPVRROGFYWRFFSHQDPDNLNDPNVQEMIDIVRW	227	
QY	284	ANLGVADILMDAVAFIWKMGVTSCEMLPCAHALIRAFNAVMRIAPAVFFKSEALVHPDQ	343	
DB	228	LGIGIDGFRLDVAPLYFEREGTTCENLPETHAFILKRVKVVDDDFPGHVLAAEAMQWEGD	287	
QY	344	WVOYIGQ-----DEQOIGYN-PIQMALLMNTLATREVNLL--LHQALTYRHHLPHTAWYN	395	
DB	288	VVEYIGDENTGDBECMAHAFPLMPRIF--MAYRRESRPISIIIAOTPPIPDMAQKGI	344	
QY	396	YVRSHDDIGMTE--ADEDAAYLGISGYDHRQFLNRFVYRRPDGSEFARGVFQYNPSTGCR	454	
DB	345	FLRNHDELTLFVMTDEBDYM-----YAEYAKDPR	374	
QY	455	VS---GTAAALVGLAODDHAHVDRIKLLYSIALSTGSLPLIYLGDEVOT-----LNDSD	505	
DB	375	MKANVGIRRLAPLLNDNR--NQIELFTALLSLSPSPVLYYDEDLGMDGVYIWGDRGG	431	
QY	506	-----WSQDSNKSDDSRMAHR--PRYNEALY-----AQRNDPSTIAAGQIYQGRH	548	
DB	432	VRIPMOWTDRNAGFSTANPGRLLYLPSPGDPVYGYQAVNVAQR--DTST--LLNFRT	487	
QY	549	MLAVQSNRPFQCGRLVTFNTNNKHLIIGYR-----NNALLAAGNSEYQGTVAHHLQ	602	

QY 617 VSINODLTQPYQVWML 634
Db 552 I-LGNETTLRPMQVWMIK 568

RESULT 8

C95854

Probable trehalose synthase protein [imported] - *Sinorhizobium meliloti* (strain 1021) me
C/Species: *Sinorhizobium meliloti*
C/Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 30-Sep-2001

C/Accession: C95854

R/Finan, T.M.; Weidner, S.; Wong, K.; Buhmester, J.; Chain, P.; Vorholter, F.J.; Hernan
Proc. Natl. Acad. Sci. U.S.A. 98 9899-9894, 2001

A/Title: The complete sequence of the 1,683-kb pSymB megaplasmid from the N2-fixing endo
A/Reference number: A95842; MUID:21396508; PMID:11481431

A/Accession: C95854

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-544 <KUR>

A/Cross-references: GB:AL591965; PIDN:CAC48499.1; PID:g15139971; GSPDB:GN03167

A/Experimental source: strain 1021, megaplasmid pSymB

R/Galbert, F.; Finan, T.M.; Long, S.R.; Publer, A.; Abola, P.; Ampe, F.; Barloy-Hubler,
Pela, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federpiet, N.A.; Fisher, R.F.;
L.; Hyman, R.W.; Jones, T.

Science 293, 668-672, 2001

A/Authors: Kaim, D.; Kaim, M.L.; Kalman, S.; Keating, D.H.; Kise, E.; Komp, C.; Lelaure,
hebaull, P.; Vandembol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yen, K.A/Title: The composite genome of the legume symbiont *Sinorhizobium meliloti*.

A/Reference number: A96039; MUID:21368234; PMID:11474104

A/Contents: annotation

C/Genetics:

A/Gene: SMD20099

A/Genome: plasmid

C/Superfamily: alpha-glucosidase; alpha-amylase core homology

Query Match 13.1%; Score 445; DB 2; Length 544;
Best Local Similarity 25.4%; Pred. No.1.3e-27;

Matches 157; Conservative 90; Mismatches 216; Indels 154; Gaps 23;

QY 84 NSSLKIDIDAR--ENNPDWILSNKQYGVYVDLFAADKLGKDKIPYQELGLTYLHM 141
Db 9 SSVYIGIDVRPFDGNGDI-----GDFIGLKERVSYLHYIGIDCVMS 52

QY 142 PLFKCEGSGDGYAVSTRDVPALGIGDLREVIALLHEAGISAVDFIFNHTSNEH 201
Db 53 PFRSP--PADNGYDVSIDYSDIPVLGTLDDFLDPLHAAGEQIRVVLDVNAHTSSEH 110

QY 202 MAQ-----RCAAGDPLFNFYIFPDRMPDQYDRTLREIFPDQ---HPGGSQLEDR 252
Db 111 WFOAARDRRCFRD-----YVWSESPP-----VAPDKTAFPGESS----- 150

QY 253 WWT-----TFNSFQMDLNSNFWFRAMAGEML-----FLANLGVDTLRMDAVA 297
Db 151 -WTVYDULAQAYFHFHFRHQPDLNTANP---AVRDELLRVVDYVLTGVDGFRVDAAP 205

QY 298 FTKWKGTCENLPQAHALIRAFNAVRIAPAVFVKSEALVHPDOVOYQITGDECOIG 357
Db 206 FYIGEGIBHADPRDQGLREMRKLVGRRODGLLGAADLAPEKLRSYFGGKIDLEF 265

QY 358 NPLQWALMNTLATREVNLLHQAITYRHNLPEHTAVNVVRSDDIGWTFADEDAAYLGI 417
Db 266 NPLCAAPASLAGQADPIGQALSTMPERPFRGANFLRNIDELNLRLEPDI----- 320

QY 418 SGYDRHQFINRPFVNFDFSGFAGVFPQYNPTGDCRVSGTAALVGLAODDPHAVRIK 477
Db 321 ---RETFAAFAPDEEMRIYGGIRRLAP-----MLDGNRA-----RLR 357

QY 478 LIXYSLSTGGLPLTYLGDVEGTAND-----DMSQDSKSDSDSRMAHPRY 524
Db 358 LAFSLTSSPVPPLIYGDIEIG-LGSDLSRPGREPVVPMQWAGSNAGPST--AQRAVL 414

QY 525 NEA-----LYAQRNDPSTAAQIYQGLHMTAVROSNRPFQGGRLVTENTNN 571
Db 525 NEA-----LYAQRNDPSTAAQIYQGLHMTAVROSNRPFQGGRLVTENTNN 571

Db 415 VQEPVTDGPFSPKRVNVEARDP-----GSLNLRVAMTLARRRHEFNRRGPFVMTLGTG 470
QY 572 KIIIGYIRNNALLAFGNESESEYPOYTAHTLQ-----AMFKADLIGKTVSL 619
Db 471 PALFA-----LAVSDGTEL--FVVLHNLSDGKRRADELPGAIDAPRNDILIGBEVEL 521

QY 620 NQD---LTLQPYQVWML 633
Db 522 SGGRLATIGLGPFGTYWL 538

RESULT 9

AF1803

Sucrose phosphorylase homolog lin2973 [imported] - *Listeria innocua* (strain Clip11262)
C/Species: *Listeria innocua*

C/Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 17-May-2002

C/Accession: AF1803

R/Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baguerro, F.; Berche, P.; Blocke
D.; Jones, L.M.; Karsc, U.

Science 294, 849-852, 2001

A/Authors: Krefz, U.; Kuhn, M.; Kunst, F.; Kurapkac, G.; Madueno, E.; Malounam, A.; M
ok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; WeilandA/Title: Comparative genomics of *Listeria* species.

A/Reference number: AB1077; MUID:21537279; PMID:11679669

A/Accession: AF1803

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-566 <GLA>

A/Cross-references: GB:AL592022; PIDN:CAC98198.1; PID:g16415513; GSPDB:GN00178

A/Experimental source: strain Clip11262

C/Genetics:

A/Gene: lin2973

C/Superfamily: alpha-glucosidase; alpha-amylase core homology

Query Match 11.5%; Score 393; DB 2; Length 566;
Best Local Similarity 24.7%; Pred. No.2.1e-23;

Matches 151; Conservative 114; Mismatches 261; Indels 86; Gaps 25;

QY 51 LMNELDSVYGN--BALLEMLEMLAQAWQSYORNSLKIDIDARENNPDWILSNKQV 108
Db 4 LKRRLSLRLSEVYVESLARIERVNGT-----KQRLVAKD-----KW--DEKDIY 48

QY 109 GVCYVDLFAADKLGKDKIPYQELGLTYL-----HMLPLFKCEGSGDGYAVSTR 161
Db 49 LITYGQFKE--SKTLPTEKMYDRYLKTFEVVHLPFYPY--SSDDGFSYIDYK 102

QY 162 DNPALGTIGDLREVIALLHEAGISAVDFIFNHTSNEHMAQRCAGDPLFNFY-IF 220
Db 103 AVNPELGDMDIKEM-----EGSARLMFPVCNHSKSEWFKRYLAGDKERQNFVEMD 157

QY 221 PDRMPDQYDRTLREIFPDQHGSGSLEDGRWVTFNSFQMDLNSNFWFRAMAGEM 280
Db 158 PD---TDLSSVTRPRATPVLTFQFASGKEG-YIWTTFSDQDLDLAFACEVLKYMIDL 213

QY 281 LPLANLGVDTLRMDAVAFIWKMGTCENLPQAHALIRAFNAVRIAPAVFVKSEA-IV 339
Db 214 MEYLBGAYVRLDAVGFMAKVPGTSSIHLDETHEIVKLFRLVMAAAGTIIITETNP 273

QY 340 HPDOVQY-IGDECOIGN-PLQWALM-----NTLATREVNLLHQAITYRHNLPE- 389
Db 274 HYDNISYFNGEKEAMVYQFPLPVLVLAIHGNAEPLSNMAKUL-----ELPBG 324

QY 390 HTAVNVYVRSDDIGWTFAD---EDAAYIGSYDRHQFINRPFVNFDFSGFAGVFP 446
Db 325 KTFPFLASHDGIIGNPVRGILPEKEILALVDLDEKEGALVSXKNPDGTS---PYEL 381

QY 447 NSTGDCRVSGTAALVGLAODDPHAVDIRIKLXSLSTGGLPLTYLGDVEGTANDDD 506
Db 382 NVTYMD-----ALSKQADTDRLSLFVLAHVAWMSIPGVPAVYVOSIIGSRNDYSG 433

QY 507 SODSNKSDSRMAHPRRYNEA-LYAQRNDPSTAAQIYQGLHMTAVROSNRPFQGGRLV 565
Db 507 SODSNKSDSRMAHPRRYNEA-LYAQRNDPSTAAQIYQGLHMTAVROSNRPFQGGRLV 565


```
Db 434 VETTG--HNSINRKYVLAETIAELNADSLRKTVDALTKLSTRKASBLFHPBLFP 490
Qy 566 TTNNTNKHIIIGYIRNN---ALIAFGNPFSEYQPTVTAHTLQAMFKAHDLIGKTVSIND 622
Db 491 EYLESTAEILFVVKRSSDASIIILNHLSE---KEVSYSLDSGVYT--NLYKSTYTGSDS 545
Qy 623 LTLQPYQVWMLE 634
Db 546 IKLRGYEFCMLK 557

RESULT 10
B85754
probable polysaccharide hydrolase yscjm [imported] - Escherichia coli (strain O157:H7, su
C/Species: Escherichia coli
C/Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 14-Sep-2001
A/Accession: B85754
R/Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glaesner, J.D.; Rose, D.J.; Mayhew
Natter, L.; Grobeck, E.J.; Davis, N.W.; Lim, A.; Dimianta, E.; Potamouzis, K.; Apodaca,
Nature 409, 529-533, 2001
A/Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A/Reference number: AB5480; MUID:21074935; PMID:11206551
A/Accession: B85754
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-434 <STO>
A/Cross-References: GB:AE005174; NID:g12515461; PIDN:AA656494.1; GSPDB:GN00145; UWGP:Z24
A/Experimental source: strain O157:H7, substrain EDL933
A/Genetics:
A/Gene: yscjm
C/Superfamily: gtfa protein

Query Match 11.3%; Score 386.5; DB 2; Length 434;
Best Local Similarity 27.6%; Pred. No. 4.6e-23;
Matches 129; Conservative 79; Mismatches 201; Indels 59; Gaps 17;

Qy 190 DTFNHTSNHENAORCAAGDPLFDFNFYIFPDRRMPDQYD-----RTLRIEFPDQHP 242
Db 3 DFCVNMKSAKSEMFKXVYLQOQPFEDFET-----AVDPQTDLAVTRPALPLLP----- 53
243 GGFSGLEDC--RWVWTFNSFQWDLNYSNPWYFRVAGEMLFANGLVILHMDVAFTM 300
54 ---FQMDNSTRLMTTTFDDQIDLNYSRSPDVLAMVDVLCLEKGAETVYLDVAGFTM 110
Qy 301 KQWGTSCENIPOAHALIRAFNAVMRLAABAVEFKSEA-IVHPDQVQY-IGQDECOIGYN 358
Db 111 KERGTSCHLEKTHLIKLRSLIIDIAQGVYIIFETVPHKONTAVFEGDEAMVYQ 170
Qy 359 PLQMALINWTLATREVNILHQAITYRHNP-ETAMVNVYVRSHDIDGWT-----FADEDA 412
Db 171 FSLPPLVLAHVQKQNVLEALCQ--WAQSLSLPSGKTWTFNFLASHDGIINPLRGLLPESBI 229
Qy 413 AVTGISGYDHRQFLNRFVNRFDGSPARQVFPQYNSTGDCRVSGTAALVGLAODDPA 472
Db 230 LALVERLQOEGALVN--WKNNDPGRS---PYEINVTYMD-----ALSRRESDEER 276
Qy 473 VDRIKLYSIALSTGGLPLIYLGEDEVGLTNDWSDSNKSDSRMAHPRY-NEALYAQ 531
Db 277 CARFLTAHAILISFPEVPAIYIGSLISHSND---YAGVEKLGYNRAINNKKTHSEITRE 333
Qy 532 KNDPSTAAQGIYQGLRHMTAVROSNPFDGGRVLTENTNNKHIIIGITRNNA-----LLAG 587
Db 334 LNDDEALTRHAYVHELISRLITLRSHNEFHPDNNFTIDTVNSSVMRLQNSNAGCNLTGLF 393
Qy 588 NFSEYQPTVTAHTLQAMFKAH-DLIGKTVSINDLTLQPYQVWMLE 634
Db 394 NVSKNIQHVNTITGLHGRDLISEVDILGNK-----ITLRPQVWMIK 434

RESULT 11
AF1416
Sturcke phosphotriase homolog lmo2735 [imported] - Listeria monocytogenes (strain EGD-e)
C/Species: Listeria monocytogenes
```

```
C/Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 27-Nov-2001
C/Accession: AF1416
R/Glaesner, P.; Frangoul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fehli, H
D.; Jones, L.M.; Karet, U
Science 294, 849-852, 2001
A/Authors: Kretz, U.; Kuhn, M.; Kunst, F.; Kurapat, G.; Madueno, E.; Maltournam, A.; M
ok, C.; Schueter, T.; Simoes, N.; Tixeret, A.; Vazquez-Boland, J.A.; Voss, H.; Wehlend
A/Title: Comparative genomics of Listeria species.
A/Reference number: AB1077; MUID:21537279; PMID:11679669
A/Accession: AF1416
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-560 <GLA>
A/Cross-References: GB:NC_003210; PIDN:CAD00948.1; PID:g16412235; GSPDB:GN00177
A/Experimental source: strain EGD-e
A/Genetics:
A/Gene: lmo2735

Query Match 10.9%; Score 373; DB 2; Length 560;
Best Local Similarity 23.2%; Pred. No. 8.2e-22;
Matches 144; Conservative 132; Mismatches 246; Indels 98; Gaps 26;

Qy 50 KLMNELD---SYGNNRSLP-MEMLLAQMSQSRNSSLKDIDARENNPWTLSN 104
Db 4 ELVNDIEVKLRKTY--QAAVQPAYEKMLACA-ENYS--NNTRGSID-----TISE 49
105 KQVGVAVVD--LFAQD---LKGDKIRPYFQELGTYLHMLPFCPCPEKSDGAYVSSY 160
50 KNVYLAAYGDSIFENKNGHPLQTLNFEQVADATDVHLPIF---PETSDDGFSVTY 106
Qy 161 RDVNPALGTIDLRVYIALHHEAGISAVVDPIFNHTSNHENAORCAAGDPLFDFNFYIF 220
Db 107 KQIDQLQDWDVQCKR-----SENERVMDLPAFANMSKSSDMPKFSDEAEYNPF--F 158
221 PDNRMPDQYDRLRIIFPDQHPGFSQLEDGFWVWTFNSFQWDLNYSNPWYFRVAGEM 280
159 IEKDSQFYNKVTY--FRTSP-LFHKYENGKELMTTSEDDDLAVNRIDCLVALITVY 214
Qy 281 LFLANGLVILHMDVAFTMKNQWGTSCENIPOAHALIRAFNAVMRLAABAVEFKSEAIVH 340
215 LFYAKQAKNTSLRDLAIGLMTSGTTCMHLEPETHIISLRLIDELVYENLQIIFETVY 274
Qy 341 PDQVQYI--GQDECOIGYNPLQMALINWTLATREVNILHQAITYRHNPETHTAVNYR 398
Db 275 HEENISYGGDGKNEAMVYQFELPPLVLAHTFCHDTYKLSKXAKSISQVSDATATFNFLEA 334
399 SHDDIGW-----TFADENAAVLGISGYDHRQFLNRFVNRFDGSPARQVFPQYNSTGDC 453
335 SHDGIQMRPATGISDEE-----INSLVQKAVQNGQVSYKDNADGTQSYELAINYGB- 388
Qy 454 RVSGTAALVGLAODDPAVDRIKL--YSIALSTGGLPLIYLGEDEVGLTNDWSDSN 511
Db 389 -----ALQNDDEDTBELVTYKIIAHSIILLTQGVPAIYHSLSKND---LVGYE 438
Qy 512 KSDSRMAHPRY--NEALYQANDPSTAAQGIYQGLRHMTAVROSNPFDGGRVLTENTNNKHIIIGITRNNA-----D 560
Db 439 ESGINRRINREKLEKQVHEKTD--TYRKTIFFSLKKLVYIRRNHAFSPFAQOEIID 496
Qy 561 GGR---LVFENTNNKHIIIGYIRNNALAFGNSEYQPTVTAHTL-QAMFKAHDLIGKT 616
Db 497 LQPDVFAIKRESBEGCTYGI-----NTSHDISKVAFSQTNLLANP 540
Qy 617 VSINODLTLQPYQVWMLE 636
Db 541 VT--SELRLTAYEVMIKKA 558

RESULT 12
AF1231
alpha, alpha-phosphotriase homolog lmo1254 [imported] - Listeria monocytogenes (strain
C/Species: Listeria monocytogenes
C/Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 14-Dec-2001
```


